2024

A SYSTEMATIC REVIEW AND META-ANALYSIS OF PSYCHOSOCIAL INTERVENTIONS FOR EARLY CHILDHOOD PROBLEMS, AGES 0-5

Hara Stephanou

A SYSTEMATIC REVIEW AND META-ANALYSIS OF PSYCHOSOCIAL INTERVENTIONS FOR EARLY CHILDHOOD PROBLEMS, AGES 0-5

A dissertation submitted in partial fulfillment
of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

to the faculty of the

DEPARTMENT OF PSYCHOLOGY

of

ST. JOHN'S COLLEGE OF LIBERAL ARTS AND SCIENCES

at

ST. JOHN'S UNIVERSITY
New York

by

Hara Stephanou

Date Submitted _____ Date Approved_____

_____ _____
Hara Stephanou Tamara Del Vecchio, PhD

ABSTRACT

A SYSTEMATIC REVIEW AND META-ANALYSIS OF PSYCHOSOCIAL INTERVENTIONS FOR EARLY CHILDHOOD PROBLEMS, AGES 0-5

Hara Stephanou

Early childhood consists of important developmental milestones, including the acquisition of daily living skills, including toileting, feeding, and sleep. While previous reviews have focused on interventions for some childhood problems, no single study has broadly assessed interventions across common presenting problems in children ages 0-5. This study systematically reviewed 41 studies on interventions for externalizing (23 studies), internalizing (3), sleep (11), feeding (3), and toileting (1) using meta-analytic methods where applicable. Overall, externalizing interventions were effective (TX1 Hedges' g = -.60; TX2 g = -.51) and largely homogeneous. Individual interventions reduced externalizing behaviors more than group or self-guided interventions (TX1 only). Internalizing studies were all randomized controlled trials aiming to reduce symptoms of anxiety and behavioral inhibition (g = -0.06, g = -.63, g = - 3.470). There was significant heterogeneity in sleep studies (TX1 g = 0.41; TX2 g = 0.46). Moderators reducing heterogeneity for sleep studies included eligibility and universality. Behaviorally-based interventions in sleep studies were more efficacious than psychoeducational interventions alone. Intervention modalities varied across feeding studies, producing small to moderate improvements (g =.13 to .69). The one toileting study found small effects favoring daytime alarms over timed potty training (g = .06). Results suggest efficacy varies by target behavior and intervention factors like format and eligibility criteria.

ACKNOWLEDGEMENTS

It is impossible to have gotten to this stage in my doctoral studies without the incredible support of so many people. Firstly, I would like to thank my advisor, Dr. Tamara Del Vecchio. Dr. Del Vecchio's patience, clarity, and expertise has been invaluable throughout my years as lab member in the CFRG and throughout this dissertation. To my committee, Dr. Lauren Moskowitz and Dr. Mark Terjesen, for their wisdom and guidance throughout my coursework, and for their support in reviewing this dissertation. Thank you especially to Genesis, Courtney, and Raul for their eagerness, patience, and skill in data coding. Special thanks goes to Dr. Celeste Sangiorgio for her meta-analytic knowledge, chats, and camaraderie as she was doing her own dissertation.

I am so fortunate to have been mentored, supervised, supported, and inspired by psychologists, healthcare and school professionals, faculty, and staff who I aspire to emulate in working with children and their families. To Dr. Dan Coletti, who had an incredibly profound impact in my pursuing the field of pediatric psychology and "on becoming a person"– thank you for believing in me from the very, very beginning. To everyone at AU's I for P, SMH, and ZHH: thank you all for igniting my excitement to pursue this dream and encouraging me right when I was applying to graduate school. An incredibly special thank you goes to Dr. Amy Nadel, Dr. Becky Lois, Dr. Lara Brodzinsky, Dr. Carly Bosacker, Dr. Mary-Beth Dewitt, Dr. Latisha Gathers-Hutchins, Dr. Victor McCarley, Dr. Yasmin Cole-Lewis, Dr. Imari-Ashley Isaksen, Dr. Beth Logan, Dr. Kevin Tsang, and the entire teams at PUFSD, Cohen's, NYU, AW, DCH, and BCH, for their supervision, encouragement, guidance, and compassion. To my graduate cohort, lab mates, and colleagues; especially Jillian, Tabitha, and Drs. Eliana, Sara,

Rahma, and Sherry – thank you for the reassurance, wisdom, humor, and support. I look up to you all so very much. I am also incredibly grateful for my SJU assistantship within the Department of Student Wellness, and to Ruth DeRosa for her flexibility and compassion as I navigated all the roles I had as a trainee. To every person I've worked with under the supervision of these excellent clinicians – thank you.

Special gratitude is reserved towards my family and friends. To my parents, who instilled a value of inquisitiveness, and have shown immense sacrifice, patience, pride, and love throughout my dream of pursuing this career. To my sister, Ki, who just "gets it," keeps the cute cat videos coming, and whom I admire so much. To my uncles, aunts, and cousins both here and in Cyprus for keeping me grounded while also lifting me up. To my grandparents and godfather for exemplifying dedication, persistence, and what it means when it "takes a village." To the Nugents and Newells (et al!) for cheering me on along the way. To all of my friends, especially Sam, Anna, Kristin, Gabriela, and Clarissa: Thank you for your love and friendship, for your understanding, and for reminding me there is a world outside of graduate school. To the kiddos in my life who I love to grow alongside with, especially Margarita, Ioanna, Elizabeth, and Danny - I adore you!

Finally, the most extraordinary thank you goes to my partner in love and in life, Daniel Nugent. From the moment we met twelve years ago you unabashedly and fiercely supported, validated, and believed in me. You are the epitome of unconditional positive regard. You deserve *all* the marshmallows. I love you, I like you, and I admire you. This would have been impossible without you.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii
LIST OF TABLES vi
LIST OF FIGURES vii
INTRODUCTION 1
Prevalence Rates 1
Gaps in the literature left by previous studies 3
Moderators 4
Parental Involvement 5
Treatment Orientation 5
Type of Control Condition 7
Intervention Delivery Format 7
Symptom Presentation and Prevention/Intervention Approach 8
Treatment Duration 10
Study Aims and Research Questions 11
METHOD 12
Study Identification and Selection 12
Inclusion/Exclusion criteria 12
Coding Procedures 14
Data Extraction for Final Included Studies 15
Analysis 17
Calculating Effect Sizes 17
RESULTS 20
Externalizing Study results 20
Study Heterogeneity and Publication bias 22
Externalizing Outcomes and Moderators 22
Externalizing Studies for Systematic Review 23
Internalizing Study results 25
Theoretical Orientation and Intervention Components 26
Study Eligibility/Symptom Severity 26
Outcome Measures/Effects 27
Sleep Study results 28
Study Heterogeneity and Publication bias 29
Sleep Outcomes 29
Significant Moderators 30
Manualized Interventions and Theoretical Orientation 30
Study Eligibility, Level of Intervention, and Intervention Type 31
Child Age 31
Non-significant Moderators 32
Sleep Studies for Systematic Review 33
Feeding study results 33
Intervention Components, Study Eligibility, and Theoretical Orientation 34
Outcome Measures/Effects 35
Toileting Study results 36
Study Description and Characteristics 36

DISCUSSION 38
Externalizing 38
Internalizing 40
Sleep 41
Feeding 43
Toileting 45
Strengths of the Current Study 46
Limitations and Future Directions 47
Implications for the Practice of School Psychology 49
Conclusion 50
Appendix A: Codebook and Moderators 52
Appendix B: Formulas used to calculate effect sizes and attrition rates 56
Appendix C: Boolean terms used for title search per topic 58
References 108

LIST OF TABLES

Table 1. Descriptive statistics – All Externalizing Studies....60
Table 2. Characteristics of Included Studies – Externalizing....61
Table 3. Effect Sizes of Included Studies – Externalizing....68
Table 4. Meta-analytic results examining categorical moderators in response to externalizing interventions (TX1)....71
Table 5. Meta-analytic results examining continuous moderators in response to externalizing interventions (TX1)....73
Table 6. Meta-analytic results examining categorical moderators in response to externalizing interventions (TX2)....74
Table 7. Meta-analytic results examining continuous moderators in response to externalizing interventions (TX2)....76
Table 8. Descriptive statistics – All internalizing studies....81
Table 9. Characteristics of Included Studies – Internalizing....82
Table 10. Effect Sizes of Included Studies – Internalizing....83
Table 11. Descriptive statistics – All sleep studies....84
Table 12. Characteristics of Included Studies – Sleep....85
Table 13. Effect Sizes of Included Studies – Sleep....90
Table 14. Meta-analytic results examining categorical moderators in response to sleep interventions (TX1)....92
Table 15. Meta-analytic results examining continuous moderators in response to sleep interventions (TX1)....94
Table 16. Meta-analytic results examining categorical moderators in response to sleep interventions (TX2)....95
Table 17. Meta-analytic results examining continuous moderators in response to sleep interventions (TX2)....97
Table 18. Descriptive statistics – All feeding studies....102
Table 19. Characteristics of Included Studies – Feeding....103
Table 20. Effect Sizes of Included Studies – Feeding....104
Table 21. Descriptive statistics – Toileting study....105
Table 22. Characteristics of Included Studies – Toileting....106
Table 23. Effect Sizes of Included Studies – Toileting....107

LIST OF FIGURES

Figure 1. PRISMA Identification and Selection of Studies Included in Systematic Review and Meta-Analysis 59
Figure 2. Forest Plot – Externalizing Interventions Included in First Meta-Analysis (TX1) 77
Figure 3. Funnel Plot – Externalizing Interventions Included in First Meta-Analysis (TX1) 78
Figure 4. Forest Plot – Externalizing Interventions Included in Second Meta-Analysis (TX2*)* 79
Figure 5. Funnel Plot – Externalizing Interventions Included in Second Meta-Analysis (TX2) 80
Figure 6. Forest Plot of Sleep Interventions Included in First Meta-Analysis (TX1) 98
Figure 7. Funnel Plot of Sleep Interventions Included in First Meta-Analysis (TX1) 99
Figure 8. Forest Plot of Sleep Interventions Included in Second Meta-Analysis (TX2) 100
Figure 9. Funnel Plot of Sleep Interventions Included in Second Meta-Analysis (TX2).....101

Introduction

The early childhood period (ages 0-5 years) is marked by the development of important milestones, including everyday skills such as toileting, feeding, and sleeping. As children develop other awareness and their own goals within such milestones, there is expected variability in behavior (Holland et al., 2017, p.10), including increases in anger, tantrums, anxiety, and aggression (Powell et al., 2006, p. 26). While milestones are developmentally transient (Briggs-Gowan, et al., 2006), how emerging skills are managed is paramount to prevent negative developmental cascades (Masten & Cicchetti, 2010) and potential psychopathology (Cicchetti & Sroufe, 2000). Most existing research has examined effects of interventions for early childhood externalizing problems. This paper provides a meta-analytic and systematic review that examines treatment modalities across externalizing, internalizing, and "daily skills" acquisition (toileting, feeding, sleeping) to determine the most effective interventions for common concerns that may arise during this developmental period.

Prevalence Rates

Approximately 8 to 10% of children under the age of five have been found to experience emotional and behavioral problems (Egger & Angold, 2006). And fifty percent of infants and toddlers who had high scores on any domain of the Infant-Toddler Social and Emotional Assessment (ITSEA) continued to have a high level of these same issues one year later (Briggs-Gowan et al., 2006). Specifically, externalizing problems tend to be the most common psychosocial concerns affecting young children. The most prevalent externalizing behavior symptoms are usually related to Attention Deficit Hyperactivity Disorder, which occurs in 2-13% of children under age 5 (Birmaher et al.,

2009; Wichstrøm et al., 2012) and Oppositional Defiant Disorder (8.3%; Birmaher et al., 2009). Conduct disorder is usually rare in very young children. However, symptoms of ODD as a precursor to CD may emerge in the preschool and kindergarten years (Holland et al., 2017).

With regards to internalizing problems, early signs of anxiety problems in young children may present as behavioral inhibition or social withdrawal. The rate of anxiety in young children to be anywhere from 1.5 - 20% depending on the types of assessments given in each study (Egger & Angold, 2006; Paulus et al., 2015; Whalen et al., 2017). The criteria for symptoms of depression between children and adults are not distinct in the DSM-5 (American Psychiatric Association, 2013; Whalen et al., 2017). According to Whalen et al. (2017), preschoolers with depressive-type symptoms typically have more anhedonia, play themes relating to death, irritable mood, and somatic symptoms. The prevalence rates for depression are low in early childhood, with estimates of 2% or less (Bufferd et al., 2011; Egger et al., 2006; Lavigne et al., 2012).

Problems that may arise during everyday skills acquisition could also be associated with behavioral and emotional difficulties. Hemmi et al. (2011) conceptualized behaviors such as difficulty sleeping and feeding as infant regulatory problems, occurring in 20% of infants (first year of life). Meltzer and Mindell (2006) note that approximately 25-40% of children exhibit some sort of sleep disturbance, typically related to having the child initiate sleep and stay asleep. Almost half (45%) of young children exhibit picky eating, fussiness, under-eating, or misbehavior during meals (Adamson et al., 2013; Morawska et al., 2014). Toilet training may also be challenging, as parents may have

incorrect expectations of when a child is aware of their physiological signals without parental oversight or reminders (Blum et al., 2003).

Gaps in the Literature Left by Previous Studies

One of the largest gaps in the existing literature is understanding what treatments work for infants and toddlers (ages 0-5) specifically. The most recent broad-scope meta-analysis conducted on youth psychosocial interventions (Weisz et al., 2017) had the youngest age of inclusion as 4 years old and the oldest age at 18. Weisz et al. (2017) stratified childhood ages to determine potential differences in effect sizes across treatments. However, nuance is still missing given the wide age range of what was considered childhood in that study (4-12 years old). A more recent review of reviews of interventions for anxiety, depression, and symptoms of ADHD (Hudson et al., 2023) in young children also only included children ages 4-9 years old. For internalizing interventions specifically, Comer et al.'s (2019) evidence-based review on psychosocial interventions for anxiety specifically in younger children still included treatments that had child participants up to 7.9 years old.

Although systematically examining interventions in younger age groups remains a gap in the literature, some topics have explored interventions with infants and toddlers more than others. For example, more meta-analyses for interventions in the externalizing behavior problem literature (e.g., meta-meta-analysis by Mingebach et al., 2018) and sleep literature (Meltzer et al., 2021a; Meltzer & Mindell, 2014; Reuter et al., 2019) have been conducted compared to internalizing problems, toileting problems, and feeding/mealtime problems. For feeding behaviors specifically, intervention research has typically focused on evaluating interventions for pediatric feeding problems inclusive of

participants who may have dependence on supplemental nutrition via feeding tubes or other oral-motor developmental delays (Lukens & Silverman, 2014), rather than age-normative mealtime difficulties (Morawska et al., 2014). Regarding daily skills acquisition, some literature has explored parenting interventions to promote early childhood development (Jeong et al., 2021), with child outcomes related to general behavior problems, attachment, or socioemotional development. Other studies exploring interventions for young children may instead look at *parental* outcomes, (e.g., decrease in maternal depression, greater responsiveness to child cues, increased self-efficacy or confidence in parenting skills (Harwood et al., 2022; Mihelic et al., 2017; Sleed et al., 2023). Additionally, more meta-analyses and systematic reviews tend to examine topics more narrowly. For example, a meta may look solely at externalizing or conduct problems or behavioral treatments not inclusive of other therapeutic modalities or theoretical orientations. One review attempted to analyze psychosocial interventions for infants and toddlers at-risk for socio-emotional difficulties on a broader scale by examining both mental health intervention *and* prevention studies with varied theoretical orientations (McLuckie et al., 2019). However, the results of the study were limited to studies occurring prior to 2012, necessitating an updated review.

Moderators

This paper aims to explore the following hypothesized moderators that may influence the efficacy of interventions for the types of problems that arise in this age period. A list of moderators and coding categories for each can also be found in Appendix A.

Parental Involvement

Young children rely on adults for activities of daily living, engaging with and learning about their world, and support with identifying and regulating emotion. This is because young children do not yet have the developmental capacity for certain cognitive tasks (Comer et al., 2019, p. 2; Kaminski & Claussen, 2017). Treatment delivery can also differ, particularly as components of certain evidence-based treatments for youth problems (e.g., CBT) may be delivered quite differently to toddlers vs. school-aged children. As a result, parents may serve as models for learning certain skills (Dasari & Knell, 2015; Eyberg et al., 2008). Parents and other caregivers may do this by incorporating changes to alter the delivery of commands to reduce child non-compliance and offering praise to increase expected behaviors. Similar logic follows with behavioral treatments for skills acquisition related to sleep, feeding, and toileting, since these skills are often first taught and managed by adult caregivers for young children. With regards to early childhood anxiety symptoms specifically, trials have shown that parental involvement as compared to treatment with the child only produced superior results (Lebowitz et al., 2020 citing Barmish & Kendall, 2005; Reynolds et al., 2012; Silverman et al., 2008). This study will explore whether parental involvement in treatment moderates the relation between a specific psychosocial intervention and resulting outcomes.

Treatment Orientation

According to a meta-analysis of psychosocial treatments for disruptive behavior problems in young children, the largest effects were associated with behavioral treatments ($g = .88$; Comer et al., 2013) with participants that were older in age and male.

In general, the largest effects were for an overall reduction of externalizing problems *(g* = .90), followed by oppositionality and noncompliance (*g* = .76) with relatively weaker effects for impulsivity and hyperactivity (*g* = .61). These results are in line with research supporting behaviorally-based parent interventions as "best practice" in treating externalizing behaviors and conduct problems in young children (Comer et al., 2013; Eyberg et al., 2008; Maughan et al., 2005).

Behavioral parent-training approaches have also demonstrated efficacy in treating common everyday issues affecting this developmental period, such as sleep difficulties (Meltzer & Mindell, 2014) and feeding problems (Lukens & Silverman, 2014). And, while there is less research in this area, behavioral parent training approaches have also demonstrated efficacy for internalizing problems (Luby et al., 2012; Luby et al., 2018). Specifically, Comer et al. (2019) evaluated treatments for anxiety and related problems specifically for young children (mean age 7.9 years). Their review included varied anxiety components, including social anxiety, behavioral inhibition, separation anxiety, and generalized anxiety. Family-based cognitive behavioral therapy (CBT) was found to be a well-established treatment, followed by Parent CBT and Group Parent CBT/Group Child CBT to be Probably Efficacious, based on guidelines set by Southam-Gerow and Prinstein (2014) for review criteria. All three aforementioned treatments include exposure-based CBT with parent involvement.

However, other treatment approaches for externalizing and internalizing behavior problems also demonstrated effect sizes in the moderate range. In a meta-analysis conducted by Lin and Bratton (2015), play-centered approaches for externalizing and internalizing behavior problems demonstrated a moderate effect size (*d* = .42 and *d* = .33,

respectively) compared to children who received an alternate intervention or no treatment at all. Although play therapy is primarily child focused, the authors also coded caregiver involvement. Those with full parental involvement had an effect size of $d = .59$ as compared to those with either partial or no involvement ($d = .33$). The authors noted, however, that these results might be confounded by the fact that parents or involved caregivers may also be "data sources," which may tend to produce better results based on their buy-in to be involved in treatment.

Type of Control Condition

One area of exploration for moderator analysis is whether intervention effects differ depending on the control condition. Weisz et al. (2017) note that this is an important consideration when critically examining effect sizes in meta-analyses, as an intervention's effect size is typically influenced by what treatment it is compared with. Meta-analyses typically lack studies directly comparing two interventions, as effects are measured against waitlist or no-treatment controls. Some research suggests that larger effects have been found for interventions that had "passive versus active control conditions" (Weisz et al., 2017). I will examine if the study design (e.g., RCT) and type of control condition used (e.g., waitlist, treatment as usual, or other active intervention) impacts treatment effect sizes.

Intervention Delivery Format

The method in which an intervention is delivered could potentially impact an intervention's effects. Some meta-analyses on parent training interventions of externalizing behavior problems (Baumel et al., 202; Tarver et al., 2014) have suggested that there is no significant difference in treatment format (self-directed vs therapist-led)

for child externalizing behavior problems. However, Lundahl et al. (2006) found individual parent training was more beneficial than group parent training. For sleep specifically, Mindell et al. (2006) noted that parents have benefited from self-guided psychoeducational sleep interventions without the need for professional guidance. To explore this moderator more specifically in this age range across presenting problems, I will explore whether individually administered, group-based, or self-guided/self-assisted interventions will differ in their effectiveness.

Symptom Presentation and Prevention/Intervention Approach

Given rapid developmental changes in early childhood, behaviors are expected to rapidly wax and wane (Holland et al., 2017). For example, younger children are developmentally expected to experience more temper tantrums than older children, Additionally, aggression typically peaks between ages two and three, and then decreases (Tremblay, 2004). While it is important to be cautious about over-pathologizing these behaviors (Holland et al., 2017), research has shown that emotional and behavioral problems that arise in early childhood can persist. As a result, understanding the efficacy of different programs that aim to either prevent or intervene on such behaviors can provide important clinical guidance.

The differences in symptom presentation in the first five years of life emphasize the importance of understanding what types of intervention and prevention programs are efficacious for infants and toddlers. Studies for this dissertation will be examined using the framework by McLuckie et al. (2019), which classified intervention mechanisms into four categories based on previous public health research: universal intervention, selective prevention, indicated prevention, direct intervention programs. Universal interventions

are programs delivered to any individual regardless of risk status or presenting problem. For example, a study examining the effectiveness of a psychoeducational sleep intervention for parents of infants seen at their pediatrician visit, despite the presence of a sleep problem, would be considered a "universal" intervention. Selective prevention approaches target children who are determined to be at risk for developing mental health disorders due to pre-determined risk factors (e.g. a program targeting behavioral inhibition in preschoolers who have parents with anxiety). Indicated prevention programs are meant for individuals that meet sub-clinical or elevated criteria that determines the potential for developing a longer-term problem (e.g., elevated scores on the ECBI, but not meeting full criteria for a diagnosis as adjusted by age). Finally, direct interventions studies target children with an established diagnosis (e.g., parent management training for children with oppositional defiant disorder).

Finally, findings across treatment studies have historically indicated that participants with more elevated symptomatology at baseline predict greater responses to treatment. For example, children with co-morbid conduct disorder and depression tend to have an increased treatment response than children with just one diagnosis (e.g., Beauchaine et al., 2005, p. 381; Beauchaine et al., 2000). However, it is important to note that, while treatment participants with more severe symptomatology appear to have large treatment responses as compared to their baseline results (as demonstrated in the MTA study; Mingebach et al., 2018; Owens et al., 2003), they may have a worse treatment response down the line. In other words, children who started off with more severe symptom scores need to go further to be considered "responders" to treatment in clinical trials, even though their overall response to treatment was substantial given their initial

symptoms. For this review, I will explore whether how children were selected for study inclusion impacts treatment effects. With regards to behaviors that the intervention addresses, I will code whether a child needed a specific symptom threshold on a cutoff score to be included in a research study (e.g., an elevated score on the Eyberg Behavior Inventory), needed a diagnosis (e.g., meets criteria for ADHD), or simply if a parent needed to state that there is a problem that they would like help with (e.g., child has difficulty with mealtime behaviors). Studies considered "universal" interventions suitable for any participant (e.g. ,study open to any parent/child between ages 3-5 visiting their doctor's office for a well visit) will also be coded.

Treatment Duration

Depending on the target problem, shorter treatments may have similar effect compared to treatments that take place over a longer period of time (Comer et al., 2013; Weisz et al., 2006). Nock (2003) suggests that there has been a push for matching clients to treatments that make the most sense for their presenting levels of severity, while also being convenient, helpful, and mindful of cost. One way to do so is to incorporate booster sessions, thus maintaining a briefer intervention but also promoting a "continued-care model" (p. 11). For example, Patterson (1974) demonstrated that a 2-hour booster session added to PMT for conduct disorder led to child behavior improvements. However, since there was no control group that did not have a booster session, results remain inconclusive (Nock, 2003, p.11). Studies will be examined to determine the impact of booster sessions on intervention outcomes. Attrition data will also be collected (see Appendix B for formulas and guidelines) to examine if drop-out rates for the intervention

as well as drop-out for intervention groups as compared to control groups moderated treatment outcome.

Study Aims and Research Questions

The purpose of this study is to update and synthesize the research on intervention addressing the psychosocial concerns (externalizing, internalizing, sleep, feeding, and toileting behaviors) for young children (aged 0-5). What is the overall effect size of treatment outcomes for these intervention practices (e.g., graduated extinction for nighttime waking) *or* packages (e.g. The Incredible Years) across externalizing studies, internalizing studies, sleep studies, feeding/mealtime studies, and toileting studies for infants/children ages 0-5? Additionally, what potential moderators explain possible heterogeneity across pooled effect sizes among these interventions or impact treatment response? Moderators proposed above are also listed in further detail in the method section below.

Method

Study Identification and Selection

I conducted a literature search using *PsychInfo* and *Medline* using the following sequel query (Boolean term): TI (child* OR infan* OR toddl* OR preschool*) AND TI (treat* OR therapy OR interven*)" to prioritize treatment studies that involved children, infants, or toddlers. Preschool was also added to ensure potential studies that categorized younger ages as "pre-school aged" to be included. Additional key words were added on the initial search term for each of the five topics (listed in Appendix C). The review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta Analyses criteria (PRISMA, Appendix D).The search to identify articles was conducted from January 27, 2020 – February 3, 2020. Studies were then downloaded into an excel database and listed by title, author, journal name, publication year, and abstract (hidden for title search). Duplicate citations were removed.

The search was limited to peer-reviewed articles written in English published between January 1, 1999 and December 31, 2019. Twenty years was used as a range for inclusion given that the Preschool Age Psychiatric Assessment (PAPA), a parent interview-based diagnostic measure for preschool ages 2-5, was first developed in 1999 (Egger et al., 2006).

Inclusion/Exclusion Criteria

The following criteria needed to also to be met for studies to be included:

- Individual or family psychosocial intervention targeting children 0-5 years of age as either *participating in treatment or primary targets of the intervention*. If there was a single participant in the study over five years old (e.g., participants ranged

from 3-7 years old) the study was not included. The definition of a psychosocial intervention was adapted from McLuckie et al., (2019) and defined as any strategy, technique, or service (be it a manualized program or components of an intervention [e.g. praise as part of behavioral parent training] that intends to "address, mediate, accommodate, affect, or reduce either the chances of onset or continues of mental health difficulties or disorders behavioral or emotional deviance, or developmental issues" (p. 3). Interventions not psychosocial in nature (e.g., music therapy, massage therapy, art therapy, animal-assisted therapy, exercise) or interventions conducted by teachers and in schools were excluded.

- Study targeted either externalizing symptoms (aggression, non-compliance, tantrums, oppositionality, conduct issues, impulsivity, hyperactivity, antisocial behaviors, disruptive behaviors), internalizing symptoms (depression, anxiety, selective mutism, behavioral inhibition), issues surrounding bedtime or sleep, feeding or mealtime problems, and toilet training or toileting concerns. Studies that included children with autism, developmental disabilities, or chronic medical conditions were not evaluated.
- Study had to target *child outcomes* and not parent outcomes of behavior. For example, if a sleep intervention study only had maternal depression outcomes and no child-related outcome measures for sleep, it was excluded.
- At least one parent-reported child symptom measure (either broadband or specific/related to the outcome studied such as duration of sleep) was included for the intervention targets listed above. If a study had multiple outcome measures, all

parent-reported measures of child behavior (such as a parent-reported Child Behavior checklist) were entered.

- If the study did not have any parent-reported measures of behavior, coders were instructed to enter any clinician reported measures (e.g., observation). For sleep studies specifically, actigraphy could be used as well.
- A control group indicating either no treatment, waitlist, treatment as usual, or another active treatment for comparison.
- At least 5 participants in each treatment condition.
- Studies need to have enough statistical information to calculate effect sizes.

A guideline of five studies per topic was set to ensure enough effect sizes were available for proper estimation of pooled effect and moderation analysis.

Coding Procedures

To ensure eligibility of studies for the title and abstract review stage, three undergraduate research assistants (RL, CH, GT) attended a virtual training during which we reviewed the code book for reviewing titles. Coders indicated a "yes, no, or unclear" in an excel spreadsheet regarding whether the article should go to abstract review. Coders were also instructed to mark whether the study title was a systematic review, scoping review, evidence-based update, or meta-analysis within our inclusion/exclusion criteria. These articles were put into a separate document and examined for any potential individual studies that should also be included. This was done in an attempt to limit bias from relying simply on titles generated from *PsychInfo* and *Medline* searches. Studies from title review marked unclear or yes had their abstracts unhidden and coded for whether the full article met criteria for final inclusion in data analysis. All titles and

abstracts were coded by the first author and a research assistant. Disagreements at the title and abstract phase were resolved by a consensus.

Data Extraction for Final Included Studies

For full article coding, two externalizing studies that did not meet inclusion criteria were picked at random and used as examples for a final virtual training meeting. All studies were coded by the first author and a research assistant. Study characteristics were coded for entry into an Excel spreadsheet, including demographics such as mean age of parents, mean age of child, parent and child gender (percentage of females), parent and child ethnicity (separately calculated as percentage of parents and children who were white), parental education (percentage of parents who had higher than a college education), and socioeconomic status. Since socioeconomic status was reported in so few studies and there was too much variation in how it was captured, it was not possible to harmonize this data for analysis. Moderators were coded as follows:

- Mean age of child was the only demographic variable explored as a continuous moderator to determine if study outcomes were impacted by infant (less than one year old)/toddler age.
- Parental Involvement was coded as yes or no. This was later modified to whether children were involved since all studies involved parents.
- Treatment orientation and other intervention characteristics were coded for treatment orientation (behaviorally based, non-behavioral, CBT, or Motivational Interviewing). For sleep specifically, interventions were categorized between psychoeducation versus implementing a specific behavioral skill.

- Type of control condition used: We coded whether studies had a passive (e.g., waitlist control) or active control group (e.g., safety intervention).
- Intervention delivery format: We coded how the intervention was delivered ((group, individual, or self-guided/self-assisted), the treatment setting(home [internet or otherwise], university, community, or primary care), and if the treatment protocol was manualized.
- Study eligibility and prevention/intervention type: We coded how a child was identified for study inclusion (needed rating scale cutoffs, parent self-referred, have age-corresponding diagnosis, no pre-identified concern, other). We also coded whether each intervention was a universal intervention, selected prevention, indicated prevention, or direct intervention.
- Treatment duration/study dosage was determined by number of study sessions and duration of these sessions in minutes. These were then multiplied to calculate a total study dosage in minutes. Studies were also coded for whether they had booster sessions.
- Study outcomes were coded by extracting the measure outcome name relevant to the study topic (e.g. ECBI for externalizing studies) and entering pre/post intervention data.
- Study attrition: pre and post participant numbers for treatment and control groups were entered to calculate study attrition rates for overall dropout and the difference in drop-out between treatment and control conditions per guidelines provided by What Works Clearinghouse (WWC) (Deke & Chiang, 2017; What Works Clearinghouse, 2014; What Works Clearinghouse, 2022). Absolute values for overall dropout and the

difference between treatment and control conditions were evaluated using a WWC attrition graph guideline (Appendix B) delineating acceptable attrition cutoffs.

Analysis

Two different analytic approaches were used for this review: meta-analysis and systematic review. Due to internalizing, feeding, and toileting topics having too few studies for moderator analysis ($k < 5$), results for these studies are presented with individual effect sizes for systematic review. For externalizing and sleep studies, random effect methods with restricted maximum likelihood models were used to estimate Hedges g pooled effect sizes and minimize Type I error (false positive). This allows for the generalizability of findings beyond studies included in the meta-analysis (Feld & Gillett, 2010). To ensure the correct directionality of effect sizes, internalizing and externalizing studies were organized so that a *negative* effect reflected greater improvement in studies, due to lower scores on scales indicating that *less* of these behaviors occurred. For toileting, sleep, and feeding studies, effect sizes were oriented so that a larger *positive* effect indicated improvements.

Calculating Effect Sizes

Study coders were instructed to enter means and standard deviations (or correlations, percentages/frequencies if those other data were not available), from each study. Since some studies provided data other than means and standard deviations), the open-source metafor package in R statistical software (Viechtbauer, 2010) was used to generate a standardized effect size (Cohen's *d*) across all study outcomes. Cohen's *d* was then recalculated into Hedge's *g* for each study. Hedges *g* is a type of effect size that is a standardized mean difference examining group differences (Field & Gillet, 2010). It is

also a variation of Cohen's *d* that corrects for bias due to sample size (Comer et al., 2013; Hedges & Olkin, 1985). The guidelines for interpreting effect size magnitude are as follows: small effect (g = .2), medium effect (g =.5) and large effect (g = .9). Once Hedges *g* was calculated for each outcome measure, effect sizes were averaged in IBM SPSS Statistics (Version 27) for each study using the aggregate function. This resulted in only one effect size per construct (e.g., general sleep outcomes). This was also done to avoid creating samples highly correlated with one another (dependent samples; Comer et al., 2013).

Moderator analysis for externalizing and sleep studies was conducted using JASP Software Version 0.16.1 (2022). Alpha levels significance was set at $p \leq .05$. For externalizing and sleep studies specifically, I tested publication bias using Egger's and Rosenthal fail-safe N tests, both in part to determine that studies with smaller effects are not missing from this dissertation. Egger's test looks at the association between effect sizes and their standard errors; "a strong association indicates small-study effects" (Shi et al., 2020). Rosenthal's fail-safe N determines how many additional studies with "zero" intervention effects are needed to raise the significance level for the entire meta-analysis. A funnel plot was generated to examine publication bias and the relationship between sample size and effect size. Funnel plots were also visually examined to see if there were any studies that deviated from the mean effect size, which may indicate publication bias (Borenstein et al., 2009; Brearly et al., 2017).

Since studies are compared to one another, I evaluated the heterogeneity between study effect sizes of studies and whether moderators affected this variability (Card, 2015). Heterogeneity of effect sizes was examined by calculating a "*Q*" statistic, which

indicates if "observed variability across effects is greater than which would be expected due to chance" (Comer et al., 2013). I then examined whether the between-group heterogeneity was significant. I^2 was calculated to determine the approximate amount of heterogeneity of effect sizes in the sample of studies (25% considered low heterogeneity, 50% considered moderate, and 75% and above considered high; Higgins et al., 2003).

Some sleep (k = 5) and externalizing (k = 7) studies had multiple intervention groups. To prevent double counting participants (Higgins et al., 2022), two datasets were created alternating different treatment groups with same controls included. These were analyzed separately, called TX1 and TX2 respectively. Attempts were made to categorize TX1 interventions as manualized interventions, those with increased professional guidance, or those with added steps to the intervention (e.g. psychoeducation plus coaching calls (TX1) vs. psychoeducation alone (TX2).

Originally, I aimed to compare studies that also had control groups that included active interventions such as educational or safety controls. However, Cochrane handbook literature suggests that different moderator effects could not be compared across subgroups without common control groups (e.g. waitlist control or no treatment, Baumel et al., 2021; Higgins et al., 2022). As a result, study control groups were dichotomized as two categories: those with active treatment controls analyzed by systematic review, and those with passive treatment conditions analyzed by meta-analysis.

Results

The initial Boolean term (Appendix C) identified 28,766 articles via *PsychInfo* and *Medline* (Appendix D for PRISMA). Results are delineated by topic in the section below.

Externalizing Study Results

The externalizing Boolean term yielded 1037 titles. Inter-rater agreement (kappa) was high for study titles to be excluded (k = .91, p =.000). Abstract review resulted in 62 studies (studies coded "no" k =.78, p =.000). Kappa calculated in Excel for moderators was moderate (75%).

Search procedures outlined in the PRISMA resulted in a final dataset of 23 eligible externalizing studies. Demographic details for the externalizing studies are presented in Appendix E, Table 1. Externalizing study characteristics can be found in Appendix E, Tables 2 and 3. Most interventions examined manualized behavioral parent training programs such as The Incredible Years and Triple P. Of the twenty-three studies, twelve studies addressed general externalizing behavior problems (Brassart et al., 2017; Brotman et al., 2003; Reid et al., 2013) including misbehavior (Bradley et al., 2003) and non-compliance (Dittman et al., 2016). "Disruptive behavior problems" was the most used term (11 studies out of 23) in describing target behaviors for externalizing interventions. Of those studies targeting disruptive behavior problems, some noted that they targeted children at the highest risk for developing conduct disorder, either due meeting elevated scores on parent measures of child behavior (Dishion et al., 2008; Hutchings et al., 2007; Posthumus et al., 2012; Sanders et al., 2000; Sourander et al., 2015), or due to having an immediate family member with conduct disorder or

Oppositional Defiant Disorder (ODD) (Brotman et al., 2003). Only one study directly addressed children meeting diagnostic criteria for ODD (Nixon et al., 2003). Two studies included children with an age-equivalent diagnosis of ADHD (Abikoff et al., 2014; Sonuga-Barke et al., 2001). Only one study was a universal intervention for parents of children interested in learning parenting skills to prevent potential misbehavior, but whose children did not have behavior problems when enrolled (Mackenzie & Hilgedick, 2000).

Seven studies had more than one intervention group. To prevent double-counting samples, studies with two intervention groups were separated into two different meta-analyses, titled TX1 and TX2. TX1 interventions were grouped by either (a) more "intensive" interventions (e.g. professional guidance versus self-directed interventions) or (b) interventions that are closer to the target age range for this meta-analysis. For example, Abikoff et al. (2014) compared the New Forest Parenting Package (NFPP) vs. Helping the Noncompliant Child (HNC) vs. waitlist control. HNC is meant to target ages 3-8 versus NFPP being eligible for use for ages 3-11, As a result, HNC was listed for TX1 meta-analysis and NFPP for TX2.

Most studies were randomized controlled trails, with two exceptions: a case control design (Posthumus et al., 2012) and a pre-test posttest comparison group design (Brassart et al., 2017). Five studies had control conditions that were not waitlist controls (e.g., educational control or treatment as usual). Three studies (Dishion et al., 2008; Shaw et al., 2006; Van Zeijl et al., 2006) only had long-term outcomes instead of immediate post-intervention data. These seven studies with either active control conditions or long-

term data only are discussed in the context of a systematic review in a separate section below.

Study Heterogeneity and Publication Bias

The meta-analysis that included the 16 externalizing studies with a passive control group indicated that TX1 (Appendix E, Tables 4, 5, and Figure 2) studies were mostly homogeneous ($Q(15)16.666$, $p = .34$; $I^2 = 22.302$). TX2 meta-analysis (Appendix E, Tables 6, 7, and Figure 3) indicated that studies were *entirely* homogeneous, thus moderator analysis was not possible TX2 ($Q(15)11.121$, $p = .74$, $I^2 = .000$). Publication bias appears to be a concern for both TX1 andTX2 based on Egger's test (TX1: $z = -2.285$, $p = .02$; TX2: $z = -1.663$, $p = .10$) and File Drawer Analysis (TX1 Rosenthal's Fail-Safe N $z = 559.000$, $p < .001$; TX2 $z = 441.00$, $p < .001$). Funnel plots for both analyses (TX1 Appendix E, Figure 4 and TX2 Appendix E, Figure 5) both appear asymmetrical, with more weight towards the left side of the effect size standard error funnel.

Externalizing Outcomes and Moderators

All 16 externalizing studies in both the TX1 and TX2 meta-analyses had measures of parent self-report of child externalizing behaviors, typically the Eyberg Behavior Inventory or Child Behavior Checklist (externalizing subscale) adapted for preschool ages. The pooled effect size indicated a moderate effect for externalizing interventions ($g = -0.60$) in the TX1 meta-analysis, with a slightly lower effect in the TX2 metanalysis ($g = -0.51$).

Given that studies were mostly homogeneous, only two moderators significantly impacted effect sizes. Study delivery method had an impact on study heterogeneity,

however only across studies in TX1 (I^2 = 0.000). Interventions that offered individual sessions showed a significant decrease in externalizing symptoms (pooled Hedges g =-0.79, p <.001). However, group interventions *increased* externalizing symptoms (pooled Hedges g =0.31, p = .01) for TX1 only. Self-assisted or self-directed interventions showed a non-significant treatment effect (pooled Hedges g = 0.17, p = .43) In terms of continuous moderators, session duration was the only significant moderator in terms of heterogeneity across effect sizes (I^2 reduced from baseline value of 22.3 to 0). The longer the intervention session, the more child externalizing behavior was endorsed by parents (pooled g = .004, p = .01). In other words, a one-unit increase (1 minute) would result in a .004 increase in externalizing behavior. The number of sessions and total dosage (sessions multiplied by study session duration in minutes), however, were both not significant.

Non-significant moderators for post-test means in the 16 studies conducted in the meta-analysis included child age, treatment setting, theoretical orientation, level of intervention/prevention, whether the treatment was manualized, child involvement, level of eligibility needed for study enrollment, presence of booster sessions, and attrition. However, this may be due to the small number of studies included in sub-group analyses. For example, there was only one "non-behavioral" study included in examining theoretical orientation, and thus a true comparison of effect sizes between non-behavioral studies and behavioral studies could not be achieved.

Externalizing Studies for Systematic Review

Seven of the 23 eligible externalizing studies had control groups or outcome reporting that was not immediately post-intervention contraindicated their inclusion in

moderator analysis. Six of these studies were RCTs with one cluster-randomized trial (Posthumus et al., 2012). Child mean ages for these seven studies ranged from 24.1 months to 50.8 months. Target behaviors addressed in the systematic review mostly addressed preventing "problem behaviors" or "disruptive behaviors," with three studies specifically addressing conduct problems (Posthumus, et al., 2012; Shaw et al., 2006; Somech & Elizur, 2012).

Most studies in the systematic review group required elevated scores on a rating scale for children to be eligible. Relatedly, all interventions for this subgroup of studies were either indicated prevention (Posthumus et al., 2012; Reid et al., 2013; Sourander et al., 2016; Van Zeijl et al., 2006) or selective prevention interventions (Dishion et al., 2008; Shaw et al., 2006; Somech & Elizur, 2012). Only one study in this subgroup required parent-reported concern regarding their children's behavior (Reid et al., 2013) without a rating scale cutoff needed for enrollment. Somech and Elizur (2012) was the only study that required a pre-school teacher's referral (teachers were not involved in the intervention, however) and elevated scores for conduct problems on the pre-k teacher version of the Strength and Difficulty Questionnaire. Interestingly, this study had the largest effect size of this specific group of studies reviewed for systematic review (g = -.43, only based on parental behavior ratings), but not the highest effect size among all externalizing studies included.

Five of the seven studies for systematic review involved manualized behavioral interventions consisting of self-directed interventions (e.g., Parenting Matters, Reid et al., 2013; Strongest Families Smart Website, Sourander et al., 2016), or group-based interventions such as The Incredible Years, (Posthumus et al., 2012), or Hitkashrut group

(Somech & Elizur, 2012). These behavioral interventions had small to moderate effect sizes suggesting reductions in child externalizing behaviors, ranging from g = -.16 to g = -0.43. Two interventions used Motivational Interview approaches as described in the "Family Check Up," which yielded small and moderate effect sizes (pooled Hedges g = -0.39 for Dishion et al., 2008; pooled Hedges g = -0.2 for Shaw et al., 2006). The one non-behavioral study in this specific subset (Video-Feedback Method; Van Zeijl et al., 2006) had the smallest effect size (g = -.09). Studies which also involved children (Shaw et al., 2006; Van Zeijl et al. 2006) had slightly higher effect sizes than the other studies for systematic review, but only if the theoretical basis for intervention was behavioral. Self-assisted/self-directed interventions had smaller effect sizes (Reid et al., 2013 g = -0.23; Sourander et al., 2016 g = -0.26) among this group.

Internalizing Study Results

A total of 1265 internalizing study titles were identified. Inter-rater agreement (kappa) was high for study titles to be excluded (k =.863, p =.000). Kappa was high for articles to be included for abstract review (k =.831, p =.000), but only moderate for studies to be excluded (k = .533, p =.000). Kappa for coding moderators/study outcomes for included studies was 83%. Full demographic variables can be found in Appendix E, Table 8. Mean child age across all studies was 48.69 months.

All studies eligible are group interventions held in university settings that targeted both behavioral inhibition and anxiety symptoms (Appendix E, Table 9). Kennedy et al. (2009) is the second efficacy trial of the Rapee et al. (2005) study. No studies were found within this dissertation's inclusion criteria that addressed depression.

Theoretical Orientation and Intervention Components

The number of sessions was similar across all three studies (6-8 sessions with a 90-minute duration). Studies consisted of two manualized treatments: the Cool Kids Program (Kennedy et al., 2009; Rapee 2005), and the Turtle Program (Chronis-Tuscano et al., 2015). Cool Little Kids is 6-week parent psychoeducation program that consists of CBT skills for parents that target their own response to their child's inhibition. The Chronis-Tuscano et al. (2015) Turtle program is an 8-week program with concurrent parent and child group treatment. Parents were taught Parent-Child Interaction Therapy (PCIT) skills modeled from PCIT for separation anxiety disorder. While parents had their groups, children also attended a brief intervention adapted from Social Skills Facilitated Play (SSFP; Coplan et al., 2010).

Study Eligibility/Symptom Severity

Studies slightly differed regarding symptom severity and existing vulnerability factors needed for study enrollment. Kennedy et al. (2009) enrolled children that had a high score on a laboratory measure of behavioral inhibition, while also requiring that at least one parent met DSM-IV criteria for an anxiety disorder. Due to this additionally required familial risk factor of a parental anxiety diagnosis, this categorized the Kennedy et al. (2009) study as a selective prevention program. The other two studies were indicated prevention programs due to children requiring a certain level of symptom severity prior to study enrollment. Participants for the Rapee et al. (2005) study required a score of 1.15 standard deviations above age-adjusted norm on the Short Temperament Scale for Children, Approach subscale *and* meet behavioral inhibition criteria on a

laboratory assessment. Chronis-Tuscano et al. (2015) required a score of 132 or higher on the parent-reported Behavioral Inhibition Questionnaire for study eligibility.

Outcome Measures/Effects

All the studies showed decreases in children's internalizing symptoms according to parent report (behavioral inhibition and anxiety; Appendix E Table 10). However, the magnitude of change in parent reports of these internalizing symptoms was much greater in the Chronis-Tuscano et al. (2015) study (pooled Hedges g = -3.47) compared to the Kennedy et al. (2009) study (pooled Hedges g = -.628) and Rapee et al. (2010) (pooled Hedges g = -.058). While there were too few studies for moderator analysis, a few differences may account for this range in effect sizes. Firstly, studies ranged as to when they collected post-study outcomes; Chronis-Tuscano et al. (2015) was the only study that reported immediate post-intervention outcomes (vs effects 6 months post-intervention). Chronis-Tuscano et al. (2015) was also the only study that directly involved children. While the Chronis-Tuscano et al. (2015) study had a suspiciously large effect size for parent reports of child anxiety/behavioral inhibition symptoms, this study also had the smallest sample size, which can impact effect size calculation. Kennedy et al. (2009) was a selective prevention study, meaning that children enrolled potentially had more vulnerability factors than the other two studies. While children did not need to have anxiety diagnosis to participate, all studies reported that most of their child participants in the treatment group met DSM-IV criteria for an anxiety disorder relative to Waitlist Control. However, Chronis-Tuscano et al. (2015) was the only study that noted that the difference in symptoms between intervention and waitlist was *not* significant. While there were too few studies to adequately determine publication bias,

the findings across these three RCTs suggest that Kennedy et al. (2009) and Rapee et al. (2010) may present stronger evidence compared to Chronis-Tuscano et al. (2015) because of their relatively larger sample sizes and lower treatment/control differential attrition rates, which could be sources of bias.

Sleep Study Results

Of the 309 titles reviewed, inter-rater agreement (kappa) was high for titles to be excluded (k=.973; p < .001). Inter-rater agreement was also high for study abstracts to be excluded (k=.807; p <.001). Inter-rater agreement calculated for moderators was moderate at 75% for all 11 eligible sleep studies that had adequate data to calculate effect sizes (see Appendix E, Tables 11, 12 and 13 for full demographics and study characteristics). The age range included infant participants immediately after birth or "zero months" to children up to 48 months; average child age was 17.93 months (*SD* =14.69).

Seven of those 11 studies (all RCTS with passive control groups) were subject to moderator analysis via meta-analysis. Mindell et al. (2009) is counted as two distinct studies, as the authors reported results separately for infant and toddler participants. Four of these studies (Mindell et al., 2011; Reid et al., 1999; Schlarb et al., 2018; Stevens et al., 2019;) had two intervention groups. To evaluate effect size differences for these other interventions without "double counting", a second moderator analysis was conducted (TX2). The remaining four studies (Adachi et al. 2009; Eckerberg, 2002; Hall et al.; 2015; Paul et al. 2016) had active control groups and are discussed in a separate systematic review section below.

Study Heterogeneity and Publication Bias

The meta-analysis that included the seven sleep studies indicated that TX1 studies (Appendix E, Tables 14 and 15 and Figure 6) were heterogeneous ($Q(7)35.023$, $p < .001$, $I^2 = 85.151$). TX2 meta-analysis (Appendix E, Tables 16 and 17 and Figure 8) yielded similar results ($Q(7)$ 41.174, $p < .0001$; $I^2 = 95.050$). However, these results are to be interpreted with caution, given so few studies. Publication bias appears to be a concern, given Egger's test results for both meta-analysis (TX1: $z = 5.263$ $p < .001$; TX2: $z = 5.170$, $p < .001$) and File Drawer Analysis (TX1 Rosenthal's Fail-Safe N $z = 118.00$, $p < .001$); TX2 $z = 82.000$, $p < .001$). Funnel plots for both analyses (TX1 Appendix E, Figure 7; TX2 Appendix E, Figure 9) indicated some asymmetry, likely due to studies having more behavioral interventions with professional guidance compared to studies with self-directed interventions.

Sleep Outcomes

The pooled effect size across sleep studies was moderate across studies across both moderator analyses (TX1 $g = .41$ 95% CI = .13, .69]); TX2 $g = 0.46$ [95% CI = -.03, .95). In terms of target behaviors for intervention, studies varied in what they considered improvements in sleep. Most studies examined whether their specific intervention improved developmentally appropriate sleep duration, reduced sleep latency (time it takes to fall asleep), or reduced night waking. Others, particularly those aimed towards toddlers, aimed to improve sleep by either trying to reduce disruptive bedtime behaviors including bedtime refusal.

For ease of moderator analysis and to further delineate intervention effects, type of sleep intervention was dichotomized as two broad categories: whether the intervention

asked parents to implement a specific behavioral technique, or if the intervention focused on providing broader parent education. Most studies included anticipatory guidance/education on typical sleep patterns for a child's age (e.g. , developing a consistent schedule, allowing appropriate time for self-soothing, etc.) However, some studies asked parents to implement specific behavioral skills. Four studies primarily used modified/graduated extinction (Eckerberg, 2002; Reid et al., 1999; Stevens et al., 2019; St James-Roberts et al., 200) compared to educational materials or waitlist. Only one study (Reid et al., 1999) had a group with a standard ignoring protocol. The Mindell et al. studies (2009, 2011) focused on examining the efficacy of a bedtime routine compared to a self-guided customized sleep profile. Schlarb et al. (2017) took a cognitive-behavioral approach, incorporating relaxation techniques for parents to model for their children.

Significant Moderators

Manualized Interventions and Theoretical Orientation. Manualized interventions appeared to have a significant effect on improving overall sleep. However, this is to be interpreted with caution as only two of the studies (Reid et al., 1999; Schlarb et al., 2017) were considered manualized (TX1 *pooled g* = 0.79, 95% CI= 0.23, 1.34, p = .01; TX2 g =1.71, 95% CI= .5, 2.32, p =.02). Studies that instructed parents to implement a *specific* behavioral technique such as a bedtime routine or modified extinction methods versus studies that gave general psychoeducation guidelines generally had higher effect sizes than those studies that solely focused on psychoeducation. Pooled effects across interventions were similar (TX1 pooled g = .48, p = .29, I^2 = 86.408; TX2 pooled g = .2, p = .11, I^2 = 94.658), however they were *not* statistically significant given the omnibus

test of coefficients. This could also be due to some interventions having very high effect size effecting how overall effects were averaged across studies.

Study Eligibility, Level of Intervention, and Intervention Type. Most sleep studies enrolled parents who self-referred to the study. Parent concerns included endorsing that their child had a sleep problem that either ranged from small to serious (Mindell et al., 2009; Mindell et al., 2011), that their child had difficulty falling or staying asleep (Stevens et al., 2019), or their child had general sleep problems (Schlarb et al., 2017).

Reid et al. (1999) was the only study in the meta-analysis requiring children to have a specific "cutoff" regarding sleep problems for inclusion. Overall, *how* parents identified sleep problems was a significant moderator. Coefficients of the model indicate that those studies that either had parents simply reporting a concern (TX1 pooled g = -1.946, $p < .001$, TX2 pooled g = -3.002, $p < .001$) or those that did not require a pre-identified sleep concern (TX1 pooled g = -2.365, $p < .001$; TX2 pooled g = -3.350, $p < .001$) had a decreased treatment effect compared to those that had to meet specific clinical cutoff set by researchers. This is also confirmed by indicated prevention studies having an increased treatment effect on improving sleep behaviors, though only for TX1 (pooled g = 0.453; $p < .001$).

Child Age. Two of the sleep studies in the meta-analysis recruited infants and their mothers immediately after birth (St. James-Roberts et al., 2001; Stremler et al., 2013). Child age was a significant continuous moderator contributing to treatment effect, with subgroup analysis demonstrating that older children generally benefited more from

sleep interventions compared to younger children (under 6 months of age) across both TX1 and TX2 groups (TX1 pooled $g = .028$, $p = .01$; TX2 pooled $g = .042$, $p = .01$).

Non-significant Moderators

Some moderator analyses could not be run due to a lack of available data, or not enough studies being part of a category. For instance, only four studies in the meta-analysis reported the number of sessions held. Most studies had a single teaching session followed by self-guided implementation, or coaching calls if there were reported difficulties. Overall, the amount of study sessions did not have a significant impact study outcomes. None of the sleep studies directly involved children in the intervention, with parents being the primary change agents. While Reid et al. (1999) and Schlarb et al. (2017) interventions did not involve children directly, there were parts of the intervention that had parents explain bedtime expectations to their children. For Schlarb et al. (2017) specifically, parents modeled CBT strategies with their children using a stuffed leopard to tell "short, calm bedtime stories" and model age-appropriate imagery and breathing techniques. Both these studies had higher effect sizes compared to all studies.

Originally, intervention format was examined as a moderator. However, all interventions in the moderator analysis were home based, either via self-administered intervention such as a customized sleep profile completed online, or directions to follow a behavioral strategy paired with telephone support. I took the additional step of coding whether interventions had a component with professional guidance to see if this impacted child sleep. There were no significant differences between interventions that received some level of professional guidance (usually telephone coaching) versus self-guided

interventions at home, thus not sufficiently explaining heterogeneity in the study outcomes.

Sleep Studies for Systematic Review

Four studies (Adachi et al., 2019; Eckerberg, 2002; Hall et al., 2015; Paul et al., 2016) were not eligible for meta-analysis due to having active control groups. These studies had some of the lowest effect sizes of all the sleep interventions. Effect sizes were likely impacted by the same moderators as studies in the meta-analysis. For example, all four studies had infants under one year old as participants. The meta-analysis also favored sleep interventions for older infants/toddlers vs infants under a year old. Effect sizes were quite low for the two studies categorized as universal interventions aimed at preventing *potential* sleep problems, (Adachi et al., 2009; Paul et al., 2016) compared to indicated prevention studies. Of note, indicated prevention studies also had an increased treatment effect with the meta-analysis group. Three of the systematic review studies were also educational in nature. Eckerberg (2002) was the only study that asked parents to implement a specific behavioral skill (graduated extinction). Consequently, Eckerberg (2002) also had the highest effect size of the four systematic review studies (g = .13).

Feeding Study Results

A total of 533 feeding studies were identified for title review. Inter-rater reliability was high for titles studies to be excluded (k = .754, p < .001). Abstract review inter-reliability was high for the studies to be included (k= .786; p <.001 but only moderate (k=*.590, p* <.001) for the full studies to be excluded. See PRISMA (Appendix D) for further details. Interrater agreement for coding moderators/study outcomes

computed in Excel was 75%. Of feeding studies examined, only three studies met criteria for inclusion. See Appendix E, Table 18 for demographic variables.

Intervention Components, Study Eligibility, and Theoretical Orientation

All three eligible feeding studies were manualized, group-based, cognitive-behavioral or behavioral approaches (Appendix E, Table 19). Morawska et al. (2014) delivered a singular two-hour, psychologist-led group modified from the Triple P Intervention for mealtime behaviors. This was the only study where parents of 37-month-old children (mean age) had to endorse that their child was experiencing mealtime or eating difficulties to participate, Group content involved setting developmental expectations for mealtimes, addressing power struggles, providing clear directions and praise for desired behaviors, and increasing consistency in providing food variety.

The Skouteris et al. (2015) authors suggested that the study was open to any parent who wanted to participate with a child aged 20-42 months at baseline, thus making this a universal intervention. This was also the longest intervention among the feeding studies, (10 weekly, 90-minute group sessions). Trained nurses and daycare workers engaged parents and their children separately in the MEND (Mind, Exercise, Nutrition – Do it!) intervention. MEND is grounded on parent-training principles covered in Triple P. Sessions provided surrounding nutrition within the five food groups and increasing physical activity, encouraging consistency surrounding mealtimes, and increasing exposures to various foods by making food preparation fun. While parents met with group leaders, children engaged in (a) guided games that encouraged physical activity, (b) “supervised creative activities” and (c) a 15-minute “healthy snack time” using

graduated exposure techniques modeled with a puppet to model trying fruits and vegetables.

Aboud et al. (2009) was coded as selected prevention study due to recruitment within a specific rural catchment area in Bangladesh with a historically documented greater risk for malnutrition. It was also the study with the youngest age group (M =13.87 months) The Aboud et al. (2009) study was unique in that community members were trained as peer coaches for parents. These coaches engaged in role plays and active problem solving using responsive feeding practices as parents fed their children. This responsive feeding intervention was a 5-session, one-month add-on to another health intervention that the parents were enrolled in located in the villages' preschools (not involving teachers). Aboud et al. (2009) was the only study that held a booster group four months after the five responsive feeding sessions.

Outcome Measures/Effects

Studies ranged both in effect size and in how they measured child mealtime and feeding behaviors (Appendix E, Table 20). Aboud et al. (2009) was the only study that did not have a parent self-reported measure of child behavior. Research coders, independent of the study intervention, used a behavioral coding system during midday meals, which measured how often a child refused food, how often a child fed themselves, and how often took a bite of food. Notably, is also the study with the highest calculated effect size. This effect size could be due to a trained observer coding behaviors, compared to parent self-report, which may rely on recall or parent attributions of child's behavior or eating patterns. Morawska et al. (2014) study was the only study that neared a moderate effect size for child feeding behaviors. This was also the only study in which

parents identified that their child had mealtime difficulties. As a result, parents may have felt that they gained more targeted skills to tailored to their child's specific mealtime problems, Skouteris et al. (2011) had the lowest effect size of the three studies (g =.131) in addressing eating behaviors such as fussiness and food neophobia. This was surprising given that children also received intervention separately from parents involving guided food exposures. However, it is unclear how often these exposures were practiced in other environments and how often these exposures occurred. Such variation in effect sizes across studies could also potentially be explained by attrition rates. Attrition across the three studies was approximately 11% (SD = 6.0%). Aboud et al.. (2009) had the least amount of attrition (4.4%), whereas Morawaksa et al. (2014) and Skouteris et al.. (2015) had similar overall dropout rates (12% and 16% respectively).

Toileting Study Results

One-hundred and eighty-two toileting studies were identified from *Medline* and *PsychInf*o. Inter-rater agreement was high (k =.88) for title exclusion, followed by (k =.79) for articles that should be excluded at the abstract level. Studies that did examine toilet training approaches were often retrospective, using questionnaires or interviews about when and how parents toilet trained their children. As a result, only one toileting study met inclusion criteria for this dissertation.

Study Description and Characteristics

The Vermandel et al. (2008) study is a randomized trial that took place in Belgium (Appendix E, Table 22). Inter-rater agreement was 100% for coding study characteristics and moderators for this article. The study was a universal intervention targeting a broad catchment area for parents with children ages who were born in 2004 or

2005, who were not yet toilet trained, and with parents who were willing to spend five consecutive days at home for the study period. Children ages 20-36 months old were randomly assigned to one of two "child-oriented" toilet training intervention approaches a daytime wetting alarm (n = 20) and timed potty training (TPT; n=19). The daytime wetting alarm was attached to the child's diaper and had a gentle ringing sound if it became wet. Once that occurred, the parent had to put the child on the potty. The TPT method consisted of scheduled visits to the toilet. For both groups, a doll was used to model toileting behaviors to children by the study investigators in the child's home. Parents were also taught to use positive reinforcement and to not use negative feedback if there was an accident, or a child refused to go to the potty. The study had a pretraining phrase, where a study investigator assessed the child at home for readiness signs of toilet training (walking/sitting down/communicate need to go to the bathroom/can pull clothes up and down), and a training phase, which was five consecutive days long. After the training phase, an investigator did a two hour observation, followed by a parental 10-hour observation one day afterwards. For this study, a child was considered toilet trained if they wore undergarments, shows awareness that they needed to void, initiated going to the toilet without prompting from their parents, and only has one leakage accident per day (Vermandel et al., 2008). This study did not consider urine/stool as separate when defining voiding/leakage. Overall, more children in the WAD-T group achieved complete dryness after the training period compared to children trained with the TPT method, despite this effect being very small (Appendix E, Table 23; g = 0.063). This could have been due to both parents and the child receiving a warning the release of urine or stool starting.

Discussion

This review aimed to synthesize literature on the efficacy of psychosocial interventions targeting every-day concerns related to externalizing, internalizing, sleeping, feeding, and toileting problems for children under age five. In examining the 41 studies across the different topics in the literature from 1999-2019, the following questions were addressed (1)What is the overall effect size of interventions for externalizing studies, internalizing studies, sleep studies, feeding/mealtime studies, and toileting studies for ages 0-5?, and (2)What potential moderators explain heterogeneity among these interventions or impact treatment response?

Externalizing

Of all the early childhood topics addressed in this dissertation, psychosocial interventions for externalizing behaviors of young children have been the most studied. Pooled effect sizes for both the TX1 and TX2 externalizing meta-analyses were moderate in decreasing child externalizing behaviors per parent report. This effect size is similar to other meta-analyses in this area (Baumel et al., 2021; Comer et al., 2013; Mingebach et al., 2018).

Due to the homogeneity of interventions across externalizing studies, moderation analysis was limited. Only two studies used non-behavioral approaches. Although research has consistently found that behavioral interventions are more efficacious than non-behavioral approaches (Comer et al., 2013), the largely homogenous sample of studies likely explains why theoretical orientation was not a significant moderator for externalizing behaviors specifically.

The TX1 meta-analytic group of individual interventions reduced externalizing behaviors only slightly more than group or self-assisted interventions, however only for the studies in the TX1 meta-analysis. Additionally, group interventions appeared to increase externalizing symptoms for TX1 only. This could be due to how the studies were coded. For example, if a study had multiple treatment delivery methods (e.g., both individual sessions and groups), I had coders what the intervention had more of. Additionally, self-assisted and self-directed interventions were grouped into one category when coding studies.

Our finding of individual sessions having slightly better outcomes, while limited to only TX1 meta-analysis, is contrary to research demonstrating no outcome differences in studies that were therapist-led vs. self-directed (Baumel et al., 2021; Tarver et al., 2014). A different meta-analysis (Harris et al., 2020) noted that for families experiencing social disadvantage, studies incorporating contact with an interventionist were more effective than those interventions that were entirely self-directed. Other studies have found that while online parent programs have reduced behavior problems, parents benefited from reminders to complete the programs (Thongseiratch et al., 2020). Future research could analyze the effect of self-directed interventions that incorporated coaching calls from those that were solely parent-guided.

The only other significant moderator among externalizing studies (though with a very minimal Hedges *g* effect) in the TX1 meta-analysis was session duration, indicating that longer sessions slightly *increased* parent-report of externalizing behaviors. It is unclear if this was for longer sessions that involved children. This is consistent with research that beyond a specific threshold, longer interventions may not have additive

effects (Bakermans-Krenenburg et al., 2003; Mihelic et al., 2017; Pinquart & Teubert, 2010). Oddly, overall dosage was not a significant moderator. This may be a due to a small number of studies, not all studies providing data on session length, and a wide range in session duration across interventions.

Internalizing

Three studies met criteria for inclusion in this dissertation, all of which were manualized programs addressing behavioral inhibition and anxiety symptoms in children with a mean age of 49 months. Studies for depression symptoms among this age group are limited, with most of this research spearheaded by Luby and colleagues (Luby et al., 2003; Luby et al., 2012). Unfortunately, these studies either (a) did not have a comparative group (Lenze et al., 2011), or (b) were excluded because they included children up to age 6 (Luby et al., 2018).

This dissertation explored *general* internalizing symptoms by aggregating measures (e.g., behavioral inhibition and anxiety together), whereas other studies/meta-analyses examined these constructs separately (Ooi et al., 2022). Literature to date (Dodd et al., 2017; Rapee & Coplan, 2010) has suggested that temperament-related characteristics of anxiety are distinct from psychopathology. This may in part explain a lower effect size from Rapee et al. (2005). While clinician ratings were not explored in this dissertation, parent ratings likely reflect more subtle changes in a child's anxiety across differing environments.

All the behavioral inhibition/anxiety interventions included parental modeling to help children practice social skills. However, only Chronis-Tuscano et al. (2015) incorporated social facilitated play for children directly. Given limited data from this one

study, it is impossible to suggest that child involvement influenced this study's high effect size. However, exploring these kinds of selective interventions and understanding the role of parent involvement is critical. As discussed in the literature review, research has suggested that parent-only interventions are efficacious for children with anxiety. Since that data has been published, discrepant data has demonstrated that there are still *no significant differences* when comparing parent-only interventions to interventions that also included children (Jewell et al., 2022; Yin et al., 2021). Of note, Jewell et al. (2022) meta-analysis excluded children described "at risk" for developing an anxiety disorder or preventative interventions. Previous literature suggests that 60% of children who have parents with anxiety disorders meet diagnostic criteria for anxiety disorders themselves (Ginsburg & Schlossberg, 2002); these same children demonstrated fewer anxiety disorders in the long-term (Bayer et al., 2018; Rapee, 2013).

Sleep

The sleep studies analyzed with meta-analytic methods indicated that sleep interventions had moderate effect sizes for both TX1 and TX2 meta-analyses. These results are similar the moderate effect sizes found in interventions for insomnia in children with a mean age of 17.6 months (Meltzer & Mindell, 2014). Multiple sleep studies had more than two intervention groups, necessitating two separate moderator analyses. However, the second meta-analysis (TX2) included a study that used extinction with a large effect size (Reid et al., 1999), thus influencing the effect among studies.

Significant moderators across sleep studies impacting the intervention effect was whether study was manualized, type of prevention/intervention study, and the child's age. Only two interventions were "manualized," (Reid et al., 1999; Schlarb et al., 2017),

which were also the two interventions with the highest effect sizes individually. However, Reid et al. (1999) had the smallest sample sizes, which may have also increased the intervention effect for both graduated extinction and standard extinction intervention groups.

Sleep studies explored in this dissertation were either indicated prevention programs or universal interventions. Universal interventions were not as effective as interventions where a sleep problem was pre-identified. Typically, universal interventions had a public-health lens geared towards younger infants to prevent sleep problems. Similarly to other studies, the mode of delivery/treatment format of sleep interventions did not have a large effect on child sleep, which was also found in the Mindell et al. (2006) review.

While effects were small, there was some reported improvement in nighttime sleep in *older* children vs. infants younger than 6 months old. This makes developmental sense, as infant sleep patterns usually stabilize at approximately six months due to maturation of the circadian rhythm (Meltzer et al., 2021b). However, these interventions mostly looked at sleep *duration* as a primary treatment outcome. There is also some disagreement in the literature about whether sleep interventions are effective for infants less than six months of age. Of all the included sleep studies, only four recruited infants under 6 months (Adachi et al., 2009; Paul et al., 2016; St. James-Roberts et al., 2001; Stremler et al., 2013). All these studies showed minimal improvements in infant sleep. While there may be some benefit to providing sleep guidance to parents of infants under 6 months old, other evaluating the benefit of doing so are still to be determined (Park et al., 2022; Reuter et al., 2019).

Feeding

Three feeding studies meeting inclusion criteria were identified, with pooled Hedges' *g* effects ranging from small to moderate. Although studies differed in their intervention components, all of them promoted modeling appropriate eating behaviors and responsive feeding for self-regulation of food consumption ("parent provides, child decides, [Satter, 1990]). Aboud et al. (2009) made note in their discussion that parents recalled few messages when it came to nutrition education specifically post-intervention. However, parents reported that they found more benefit from direct modeling on how to react during mealtime behaviors (e.g., food refusal). Aboud et al.'s (2009) responsive feeding intervention was an add-on to a larger intervention educating parents on "gentle discipline." This may have given parents some additional skills in the potential overlap of misbehaviors that may also occur during mealtimes/preventing further coercive cycles. While anecdotal from only one study, this aligns with strategies used in behavioral parent training that have long demonstrated efficacy in the externalizing literature, including strategies used in the Hassle Free Mealtimes Triple P study (Morawska et al., 2014).

Despite the direct involvement of children in the Skouteris et al. (2015) intervention, this intervention had the lowest effect size compared to the other feeding studies. It is important to note that child age may be playing the role, as the children in the Aboud et al. (2014) study were much younger (13.87 months on average) compared to the Skouteris et al. (2015) study (33 months). Developmentally, this is one of the prime times during which "picky eating" occurs. Carruth et al. (2004) noted that the percentage of children identified as "picky eaters" is approximately 50% at age two. In the Skouteris et al. (2015) study discussion, authors noted that food neophobia was improved at *12*

months instead of post-intervention. This could be a result of continued exposure to new foods, which is supported by research suggesting an average of 15 exposures of a new food are needed before it is accepted by a child (Wardle et al., 2005).

Overall, very few feeding studies were eligible for this dissertation. Firstly, there does not seem to be a universally accepted definition of "selective eating" or mealtime problems outside of more severe eating problems (e.g. Avoidant Restrictive Food Intake Disorder; Tanner et al., 2022; Taylor et al., 2015). Another area of literature regarding feeding/increasing diversity of foods presented to young children is within the obesity *prevention* literature, which are often larger public health initiatives (typically selective prevention or universal interventions). For this dissertation specifically, coders were told that obesity *intervention studies* were a "medical exclusion" criteria. However, obesity *prevention* study abstracts were examined to see if they utilized psychosocial frameworks/interventions also appropriate for mealtime problems/selective eating. In reviewing abstracts, these studies were often multi-modal universal intervention programs, and attempted to also target parenting behaviors related to sleep, amount of screentime allotted, and physical activity. For the studies that did incorporate a child-behavior related measure regarding mealtime experiences, there was often an issue of establishing a "baseline" for effect size calculation (Daniels et al., 2009; Helle et al., 2017; Savage et al., 2018). For example, The Children's Eating Behavior Questionnaire is normed for children ages two and older, whereas some of these studies began during the introduction of solid foods (6 months). Given that child feeding is also heavily developmentally dependent (e.g. bottle feeding, breast feeding) it is difficult to compare feeding patterns when breast/feeding or formula feeding compared to introducing solid

foods during a longitudinal study period. Finally, different keywords could have identified additional interventions regarding exposures to newer foods more generally instead of selective eating patterns or mealtime behaviors. These include flavor conditioning, associative conditioning, flavor-flavor learning, and fruit/vegetable acceptance.

Toileting

Only one study from Boolean terms/literature searches met inclusion criteria (Vermandel et al., 2008). More children in the auditory daytime alarm group achieved dryness at the end of the five-day intervention compared to children trained with the timed potty training method, despite this effect being quite weak.

Both approaches in the Vermandel et al. (2008) study emulate the original Azrin et al. (1974) approach to toilet training, with slight variations. In their discussion section, Vermandel et al. (2008) noted that parents reported more oppositionality and behavioral concerns with the timed potty training (TPT) approach. This may be due to the TP-T approach incorporating scheduled toilet visits at 2–3-hour intervals, with no diaper during the day. As a result, there may have been an increase the number of demands placed on children versus an alarm sounding only when a diaper was moist. Other research has also indicated that more "intensive" toilet training that requires parents asking a toddler to use the toilet more than three times a day *prior to 27 months* may be associated with stool withholding or toileting refusal (Blum et al., 2003). Of note, the average age for children in the Vermandel et al. (2008) study was 26.5 months.

In reviewing the toileting literature outside of the one study that met criteria for this dissertation, most studies explored structured behavioral approaches to toilet training.

However, virtually no studies compared different toilet training approaches to each other each other, as toilet-training *actively* took place, often using either case-control, cross-sectional, observational, longitudinal, or prospective designs, most without comparative groups. Definitions of "toilet training" are also inconsistent (e.g., daytime vs. nighttime control, bladder vs. bowel control, threshold for accidents, etc.). Additionally, "structured behavioral approaches" are studied by examining *variations* of the approach, such as using a wetting alarm, intensive toilet training regardless of readiness (negative reinforcements of accidents), assisted infant toilet training, and elimination communication (de Carvalho Mrad et al., 2021).

Strengths of the current study

This dissertation is the first study I am aware of to attempt to systematically review and meta-analyze data by calculating effect sizes across interventions for multiple, common, every-day problems for children exclusively under five years old. I attempted to cast a broad net to include both intervention and prevention studies across topics. Additionally, an attempt was made to distinguish preventative vs. intervention studies as a moderator of effect size. While studies comparing different active interventions were both few and not eligible for meta-analytic methods, they were discussed in the content of systematic review. Multiple databases and search terms were used, in addition to reviewing systematic reviews, meta-analyses, and evidence-based reviews for additional eligible studies. Coders not only extracted data for analysis, but provided important inter-rater reliability for data extraction.

Limitations and future directions

Although I aimed for breadth in generating Boolean terms and searching within other meta-analyses and reviews, several limitations arose. Firstly, some studies were not included for review based on exclusion criteria more generally (e.g. single-subject designs, dissertations and unpublished papers). Additionally, studies had to be published in English, which may have excluded (a) research done in other countries besides English-speaking nations and (b) culturally-adapted interventions. This dissertation also only explored interventions for neurotypical and typically developing youth. As such, these results may not speak to what may best serve other populations. In order to be more inclusive of different populations and interventions, future research should incorporate varied study designs and settings.

In conducting meta-analysis, both sleep and externalizing studies demonstrated indications of publication bias based on Egger and Fail-safe n tests. While many studies were RCTs, several studies in this dissertation had small samples, which likely impacted effect size calculation. Additionally, a small number of studies did not allow for advanced meta-analytic methods such as meta-regression to determine correlations among moderator groups. Data provided by studies across topics to calculate effect sizes was mixed. Occasionally, data other than means and standard deviations had to be used (e.g., confidence intervals) to generate an effect size. While I used whatever data was available that could adequately be converted to Hedges *g*, contacting authors may have provided additional information needed to calculate effect sizes and potentially target an even greater breadth of studies.

A common theme in the literature is *defining* constructs for behaviors that children exhibit (e.g., when is toilet training actually complete?) and how constructs should be measured. For example, the definition of "risk" can especially be further delineated for selective prevention programs, as there is overlap with children who are "indicated" for an intervention. For example, having a high ECBI score could imply potential "risk" of developing more severe externalizing problems. However, interventions targeting externalizing symptoms of children who have parents with elevated scores for depression and family stressors is also deemed "at risk" in the literature (McLuckie et al., 2019). Further research is needed to determine a more "probabilistic risk (e.g., belonging to a target group at risk)" versus those children have who meet "subclinical" criteria for something like externalizing behavior problems (McLuckie et al., 2019, p. 11).

Defining an intervention's theoretical orientation could have also been further delineated. I initially thought to code interventions as "behavioral" or "non-behavioral" when first proposing this dissertation. However, once establishing coding criteria, I found that it was important to add nuance to whether the interventions were more psychoeducational, behavioral skill-based, or had cognitive components for parents. It will likely be beneficial for future studies to explore such nuances, as well as specific components of an intervention that may contribute to efficacy (e.g., psychoeducation vs. asking parents to try a specific behavioral skill).

Effect sizes were aggregated to provide one effect size per the related topic of the study, which is helpful in understanding general outcomes. However, this attempt for breadth takes away from understanding the reduction of *specific* types of behaviors or outcomes that could be worth analyzing. Potentially eligible studies were also excluded

due to not providing adequate baseline data. This could also be a result of validated parent-self report measures being limited, particularly for such young ages, or needing alternate ways to measure potential problems.

This dissertation only explored parent reported changes of *child* behavior. A future direction could be to examine *parental* or caregiver changes in behavior as a result of the intervention, however not all studies reported such measures. Looking at parent behavior change outcomes can contribute to understanding understand what behaviors parents may be more likely to change (e.g. increasing parental "cry tolerance" to implement specific bedtime routines; Kahn et al., 2020). Exploring outcomes of child behavior from different raters could also add insight towards change across environments with differing demands (e.g. feeding at daycare vs home).

For feeding and sleep behaviors specifically, there are more "population based" or "public health" studies compared to externalizing and internalizing behaviors; only one externalizing study was identified as a universal intervention. Additionally, trials examining universal interventions are still typically evaluated using clinical rating scale norms instead of population effects (Bayer et al., 2010; Sarkadi et al., 2014). Population effects may detect smaller differences that account for large differences or impact across a population (Bayer et al., 2010; Sarkadi et al., 2014). As discussed in Weber et al. (2019), lower initial problem intensity can affect effect size magnitude, and thus may not demonstrate as much symptom change (Reyno & McGrath, 2006).

Implications for the Practice of School Psychology

Studies that took place in school settings or with day-care providers were excluded from this analysis. However, this provides a future opportunity to evaluate such

related interventions with school providers. Professionals in these settings can serve as an important access point to identifying and screening for potential problems. They can also provide important collateral observations across multiple raters on child behaviors. Other studies (e.g., MEND trial; Skouteris et al., 2016) have noted potential cost benefits, increased reach, and sustainability in implementing interventions with fidelity from reliable professionals in schools. School providers can provide consistency in offering tiered prevention and intervention programs, consulting with parents and other professions, and progress monitoring throughout a child's developmental trajectory. As mentioned in Holland et al. (2017), the multi-tiered system of support can be helpful in how school psychologists can target prevention and intervention efforts for students. In terms of positive behavioral interventions and supports (PBIS) in preschool or kindergarten classrooms, embedded tiers of universal prevention programs that target a whole population/classroom fit well with the public health model also discussed in McLuckie et al., (2019). School psychologists can aid teachers in managing child classroom behaviors related to both early externalizing/internalizing behavior problem risk, as well as associated daily skills acquisition. School psychologists can also be a valuable resource for community outreach, including providing psychoeducation and brief skill-based suggestions at parent-teacher conferences, liaising with local pediatricians and early childcare providers, and establishing resource banks in community settings where children and families gather.

Conclusion

Overall, findings from this dissertation illustrate the importance of critically evaluating interventions that address every-day problems and risk factors in the early childhood

years. It is important to consider how participants are eligible for study entry/level of symptom severity at baseline as well as the nature of preventative/intervention programs offered to children and families, particularly for sleep studies. For externalizing studies specifically, individual interventions may be more efficacious, however it is still unclear if individual interventions with shorter session duration apply across different symptom severities and risk factors. Few studies were found for internalizing, toileting, and feeding for this young age group. However, this dissertation emphasizes the importance of exploring such interventions not just in clinical frameworks but determining efficacy for broader population effects.

Appendix A

Codebook and Moderators

Codebook (*=indicated moderator)

Identifiers

Meta-analysis study ID (STUDYID) – paste from final study list

Title of the study (STUDYTITLE) – paste from article

All study authors (STUDYAUTH) – paste from article

Publication year (YEARPUB) – paste from article

Country study took place (COUNTRY) – paste from article

Study Characteristics:

Study topic (STUDYTOPIC)
- Externalizing
- Sleep
- Internalizing
- Feeding
- Toileting

Study design (STUDYDESIGN)
- Randomized Controlled Trial
- Other (Write in)

What type of control group was used? (CONTROL_TYPE)
- Treatment as usual
- Waitlist Control
- No treatment
- Educational control
- Other (write in)

Total sample size – paste from article (TOTAL_N)

How many treatment (intervention) groups are there? – count and write in (ACTIVEK)

*Treatment Setting – Where was the majority of the study held? (TXSET)
- Home based
    - For sleep, indicate if home based is internet or other
- University Based

- Community Based (e.g. community clinic)
- Primary Care or other health setting
- Other (write in)

*Study Delivery Format
- Individual sessions
- Group
- Self-assisted or Self-guided/directed

*Treatment theoretical Orientation (THEOORIENT)
- Non behavioral (e.g. play-therapy, psychodynamic/ attachment-based, supportive counseling)
- Behaviorally Based Treatment (e.g. operant or respondent conditioning, social learning theory, modeling, most behavioral parent training/ positive reinforcement, effective limit setting, problem solving skills, social skills training etc.)_
    - For sleep, specifically code:
        - psychoeducation only
        - Behavioral skill implementation (e.g. graduated extinction, specific bedtime routine)
- Other (write in)
    - e.g. Motivational Interviewing, CBT

***What was the level of intervention/prevention and the population targeted** (per McLuckie et al., 2019) **(INT_PREV)**

- Universal intervention- programs are offered to the broadest range of infants/preschoolers/families
- Selected prevention - intervention is for a "high risk" group (e.g. for children of parents with mental health concerns)
- Indicated prevention – provided to children with no formal diagnosis but have subclinical problems, difficulties, or elevated scores on screening tools
- Direct interventions – Direct psychosocial interventions for children with an established age-appropriate diagnosis

*Was the treatment manualized (should state this in text, usually is part of a treatment package like Incredible Years or PCIT)? (MANUALIZED)
- No
- Yes

*Child Involvement (WEREKIDSINVOLVED)?
- No
- Yes

Inclusion

*How is a child identified for inclusion in a study? (ELIGIBILITY)

- RS: Had to meet/be above a clinical cutoff on a scale when screening for study inclusion
- PC: parent reported concern or self-referred (no rating scale cutoff or screening cutoff needed)
- DX: had to have a diagnosis to be included in the study
- NC: no parent concern about behaviors/symptoms or no diagnosis needed to be included in the study (likely universal intervention
- O; Other (write in)

Session Information: If study does not have specific information, leave blank.

*Number of Treatment Sessions– write in number

*Session Duration (for each individual session) in minutes - write in number

*Dosage in minutes (DOSAGE_REPORTED_TX1) – will be calculated by multiplying session number by session duration

*Was there a follow-up or booster session? (BOOSTERTX1)
1 = No
2 = Yes

*Dropout rate Overall (Overall_Dropout) - calculated with What Works Clearinghouse (WWC) formula in excel (formula in Appendix B)

*Dropout rate TX/CTRL Differential (TX_CTRL_Diff_Dropout) – calculated with What Works Clearinghouse (WWC) formula in excel (formula in Appendix B)

*Was attrition deemed acceptable based on WWC standards (categorized with What Works Clearinghouse graph, Appendix B)

Demographics/Sample Characteristics: leave blank if no data. If possible, write in number for *whole study* rather than treatment and control group separately.

Mean age of Parents (MEANAGEPAR_OVERALL) – write in

*Mean age of child (MEANAGECHILD_OVERALL) – write in

Parent Percent Female (PER_FEMALE) – write in

Parent Percent White (PER_WHITEPARENT) – write in

Child Percent Female (PER_FEMALECHILD) – write in

Child Percent White (PER_WHITECHILD) – write in

Parent Percent had college degree (PER_COLLEGE) – write in

Outcomes – repeat for each outcome measure, group, and time delivered (e.g. pre-, post-)

Treatment Outcome Measure Name – Paste Name of rating scale directly from the study (e.g. ECBI) (TX1_PREOUTCOME1_NAME)

Sample size (n) for treatment group at baseline (TX1_PREOUTCOME_N)

Mean score for treatment group at baseline (TX1_PREOUTCOME1_MEAN)

Standard deviation for treatment group at baseline (TX1_OUTCOME1_SD)

Treatment Outcome Measure Name – Paste Name of rating scale directly from the study (e.g. ECBI) (TX1_POSTOUTCOME1_NAME)

Sample size (n) for treatment group at post-intervention - TX1_POSTOUTCOME_N)

Mean score for treatment group at post-intervention (TX1_POSTOUTCOME1_MEAN)

Standard deviation for treatment group at post-intervention (TX1_POSTOUTCOME1_SD)

Notes for Outcomes:
*If outcome measures do not have a mean and standard deviation, write in scores provided (e.g. odds ratio, confidence intervals, percentages, etc) and associated data
*If data is missing, try to check text of study or potential supplemental materials (not just tables) to see if can find data there

Appendix B

Formulas used to calculate effect sizes and attrition rates

Due to studies not always having means and standard deviations for effect size calculations, some different formulas were used to be able to calculate effect sizes. Difference in reporting includes studies that reported frequencies and percentages for categorical outcomes, means and standard errors or 95% confidence intervals instead of standard deviations, and odds ratios. The formulas were pasted into excel. Once this was done, the excel sheet was transferred into R to convert data to Cohen's *d*, and then ultimately to Hedges *g*.

Formula for Cohen's *d* (from means, *SD*, and *n*):

$$d_{PPWC} = \frac{(T_{post} - T_{pre}) - (C_{post} - C_{pre})}{S_{pre}},$$

$$where\ S_{pre}\ is\ the\ pooled\ SD\ for\ the\ two\ groups\ at\ baseline,$$

$$S_{pre} = \sqrt{\frac{(n_t - 1)s_t^2 + (n_c - 1)s_c^t}{n_t + n_c}}$$

- d_{PPWC}Formula (pre-post with control designs) calculates standardized mean differences for continuous outcomes with M, sd, and n for pre-post designs (Hoyt & Del Re (2017); formula from Carlson and Schmitt (1999)
- Formula for pooled sd: Borenstein et al. (2009); Card (2016, pg. 124); Thalheimer and Cook (2002).

Formula to calculate Odds Ratios (OR) from Frequencies:

	Event	*No event*	
Treatment	*a*	*b*	n_{treat}
Control	*c*	*d*	$n_{control}$
	n_E	n_{noE}	

$$OR = \frac{a/b}{c/d}$$

- Note: pre is no event, post-data is event
- Formula references: Card (2016, pg. 95) and Harrer et al. (2021).

Formula to calculate standard deviations from standard errors/95% CI

$$SD = \sqrt{N * (Upper\ limit - Lower\ limit)/3.92}$$

-Formula from Higgins and Green (2011)

Formula to convert r correlations into standardized mean differences (d); Borenstein et al., 2009):

$$d = \frac{2 * r}{\sqrt{1 - r^2}}$$

Conversion of d to Hedge's g: Turner and Bernard (2006):

$$Hedge's\ g = d * (1 - \frac{3}{4(n_I + n_c) - 9})$$

Aggregating multiple outcomes within a study:

If studies reported multiple outcomes, they were averaged together within a study to result in one effect size. This approach is suggested by Turner and Bernard (2006).

Calculating Attrition Rates and Thresholds from What Works Clearinghouse (2014; 2022; Deke & Chiang, 2017) and Institute of Education Sciences (n.d.) Attrition. [PowerPoint Slides]. https://ies.ed.gov/ncee/wwc/Docs/OnlineTraining/wwc_training_m2.pdf

The formula for calculating Overall Dropout rates:

$$Overall_{Drop} = \frac{((Pre\ N_{TX} + Pre\ N_{CTRL}) - (Post\ N_{TX} + Post\ N_{CTRL}))}{(Pre\ N_{TX} + Pre\ N_{CTRL})}$$

The formula for calculating Treatment vs Control Differential Dropout rates:

$$TX\ vs\ CTRL\ Diff_{Drop} = \frac{(Pre\ N_{TX} + Post\ N_{TX})}{(Pre\ N_{TX})} - \frac{(Pre\ N_{CTRL} + Post\ N_{CTRL})}{(Pre\ N_{CTRL})}$$

Absolute values of the two calculated formulas above are used to determine thresholds on graph provided by document link above.

Appendix C

Boolean terms used for title search per topic

Initial query: TI (child* OR infan* OR toddl* OR preschool*) AND TI (treat* OR therapy OR interven

The following specific terms were added for each topic:

Externalizing: (TI (child* OR infan* OR toddl* OR preschool*)) AND (TI (treat* OR therap* OR interven*)) AND (TI (external* OR opposition* OR defia* OR hyper* OR impul* OR inatt* OR disrup* OR aggress*))

Sleep: ((TI (child* OR infan* OR toddl* OR preschool*)) AND (TI (treat* OR therap* OR
interven*)) AND (TI (slee* OR bed* OR insom*))

Internalizing: (TI (child* OR infan* OR toddl* OR preschool*)) AND (TI (treat* OR therap*
OR interven*)) AND (TI (intern* or anx* or dep* or inhib* or withdr* or fear))

Feeding: (TI (child* OR infan* OR toddl* OR preschool*)) AND (TI (treat* OR therap* OR
interven*)) AND (TI (feed* OR meal* OR food* OR eat*)

Toileting: (TI (child* OR infan* OR toddl* OR preschool*)) AND (TI (treat* OR therap* OR interven*)) (TI (toilet training OR potty training OR Toileting)

Figure 1

PRISMA Identification and Selection of Studies Included in Systematic Review and Meta-Analysis

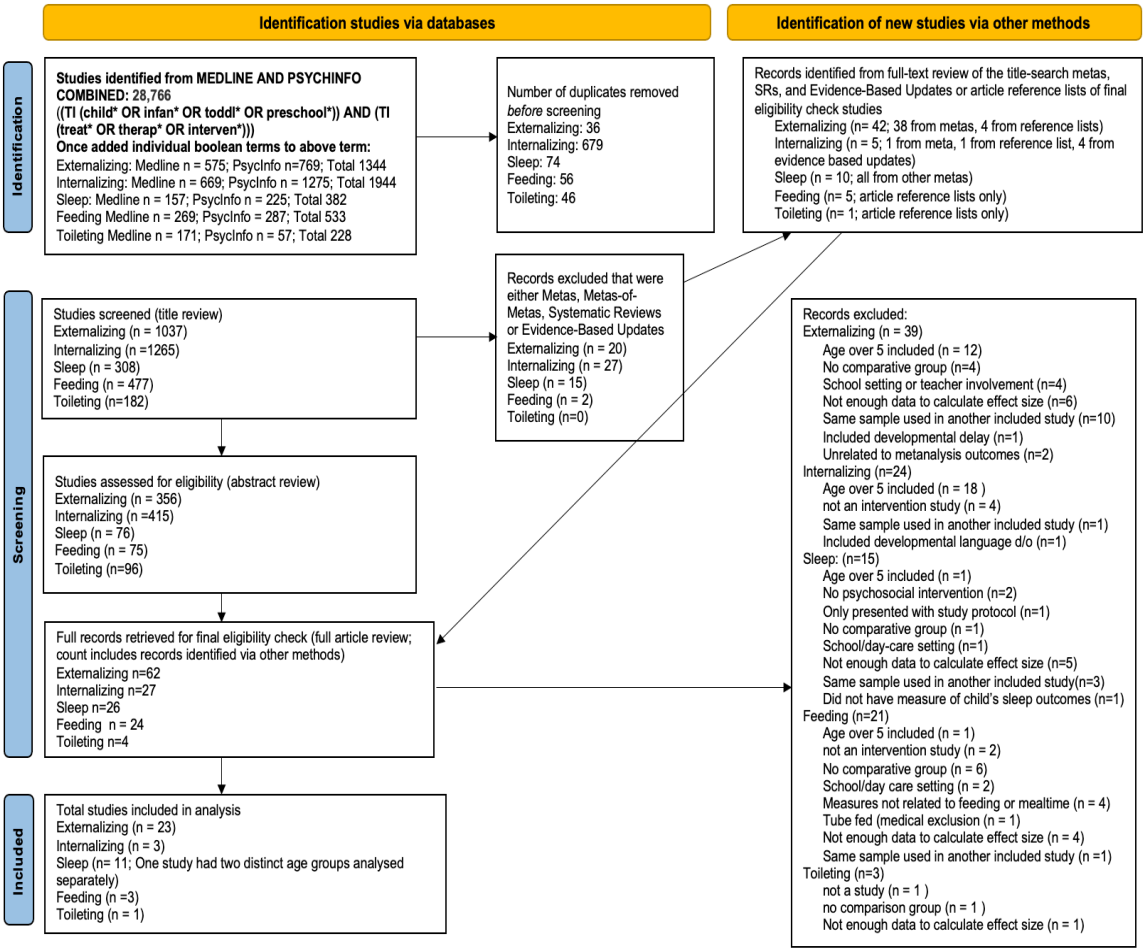

Table 1

Descriptive statistics – All externalizing studies

	k	*Mean*	*SD*	Range
Parent age (yrs.)	14	33.923	4.188	20.950 - 37.365
Child age (mos.)	20	42.820	7.059	26.100 - 51.200
Percent female parent	13	0.918	0.059	.800 - 1.000
Percent white parent	6	0.640	0.304	.150 - .970
Percent $\geq$ college education	16	0.552	0.265	0 - .931
Percent female child	23	0.352	0.119	0 - .508
Overall attrition	18	.1353	.0734	.0207 - .3175
Treatment vs. control differential attrition	18	-.0324	.1442	-.3056 - .3455

Table *2*

Characteristics of Included Studies – Externalizing

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
^Abikoff et al., 2014 United States	RCT; WLC	164 HNC: 63 NFPP: 67	42.84 mos.	Dx	ADHD externalizing behaviors (oppositionality /aggression/ defiance)	Helping the Noncompliant Child (HNC)+ New Forest Parenting Package (NFPP)+	HNC: 8/60/480 OD:4.1% TCD: -6.3% NFPP: 8/90/720 OD: 5.9% TCD: -9.0%	HNC: B/DI/C/I/CO/NB NFPP: B/DI/C/I/HB/NB
^Bradley et al., 2003 Canada	RCT; WLC	198 (89)	45.6 mos.	PC	Child misbehavior	1,2,3 Magic parenting group+	3/120/360 OD: 12.1% TCD: 5.7%	B/IP/NC/G/CO/B
^Brassart et al., 2017 Belgium	pre/post comp.; WLC	41 (21)	50.56 mos.	RS	Externalizing behavior problems	Parent-Implemented Verbal Responsive Intervention	8/90/720 OD: NP TCD: NP	NB/IP/NC/G/CO/NB
^Brotman et al., 2003 United States	RCT; No tx	30 (16)	44 mos.	O	Externalizing behavior problems	Multicomponent (CBT & IY)+	60/90/5400 OD: 6.7% TCD: 14.3%	CBT/SP/C/G/CO/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
^Dittman et al., 2016 New Zealand and Australia	RCT; WLC	85 (45)	43.5 mos.	PC	Noncompliant behavior problems	Triple P Dealing with Disobedience Discussion Group+	1/120/120 OD: 21.2% TCD: -30.6%	B/IP/NC/I/U/NB
^Hutchings et al., 2007 Wales/UK	RCT; WLC	153 (104)	46.3 mos.	RS	Antisocial, hyperactive, and disruptive behaviors	The Incredible Years+	12/150/1800 OD:13.1% TCD:-13.2%	B/SP/NC/G/CO/NB
^Mackenzie and Hilgedick, 2000 United States	RCT; no tx	46 CAPP= 16 Book= 15	51.2 mos.	NC	Externalizing behavior problems	Computer-Assisted Parenting Program (CAPP) Booklet group	CAPP: 4/90/360 OD: NP TCD: NP Booklet: NA due to SA OD: NP TCD: NP	CAPP: B/UI/NC/SA/CO/B Booklet: B/UI/NC/SA/HB/B
^Markie- Dadds and Sanders, 2006 Australia	RCT; WLC	63 (32)	43.09 mos.	RS	Disruptive behavior problems	Self-Directed Triple P+	NP due to SA OD: 31.7% TCD: -5.3%	B/IP/NC/SA/HB/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
^Morawska et al. 2011, Australia	RCT; WLC	67 (23)	43.56 mos.	RS	Disruptive behavior problems	Triple P Discussion Group+	1/120/120 OD: 20.9% TCD: -5.3%	B/IP/NC/I/U/NB
^Morawska & Sanders, 2006 Australia	RCT; WLC	126 TASD-BFI: 43 SD-BFI: 42	26.1 mos.	PC	Disruptive behavior problems	Self-administered Triple P plus brief therapist telephone assistance (TASD-BFI)+ Adapted self-administered Triple P (SD-BFI)+	Both tx lasted 10 weeks TASD-BFI: 10/10/100 OD: 7.1% TCD: 5.1% SD-BFI: NA due to SA OD: 14.5% TCD:-9.3%	TASD-BFI: B/IP/NC/SA/HB/NB TASD-BFI had additional weekly telephone consultations initiated by a psychologist, max. call was 30 min Coded as SA due to that being the majority of the intervention. SD-BFI: B/IP/NC/SA/HB/NB
^Nicholson et al., 2002 United States	RCT; WLC	26 (13)	12-60 months	RS	Externalizing behavior problems	STAR Parenting Program +	10/90/900 OD: NP TCD: NP	CBT/SP/NC/I/U/NB
^Nixon et al., 2001 United States	RCT; WLC	34 (17)	46.64 mos.	Dx	Disruptive behavior problems and hyperactivity	PCIT+	12/60/720 OD: NP TCD: NP	B/DI/C/I/CO/B

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
^Nixon et al., 2003 United States	RCT; WLC	54 STD: 17 ABB: 20	46.75 mos.	Dx	Oppositionality/ Defiance and disruptive behavior problems	Standard PCIT (STD)+ Abbreviated PCIT (ABB)+	STD: 12 sessions/ 930 min total OD: 15.0% TCD: -17.2% ABB: 12 sessions/ 570 min total OD: 9.8% TCD: -7.5%	STD: B/DI/C/I/CO/B ABB: B/DI/C/I/CO/B Still had individual intervention, however five face-to-face sessions replaced w/ 30-minute calls and video lessons
^Perrin et al., 2014 United States	RCT; WLC	273 R-IY: 89 NR-IY: 123	36.36 mos.	RS	Disruptive behavior problems	Randomized Abbreviated Incredible Years protocol (R-IY)+ Non-Randomized Abbreviated IY (NR-IY)+	Both groups: 10/120/1200 R-IY OD: 18.7% TCD: -1.1% NR-IY OD: 33.7% TCD: -23.4%	Both groups: B/IP/NC/G/CO*/NB *recruited from primary care

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
^Sanders et al., 2000 Australia	RCT; WLC	305 EBFI: 76 SBFI: 77	41.12 mos.	RS	Disruptive behavior problems	Triple P Enhanced Behavioral Family Intervention (EBFI)+ Triple P Standard Behavioral Family Intervention (SBFI)+ (Triple P self-directed BFI not examined for this dissertation)+	EBFI: 12/60/720 OD: 15.7% TCD: -15.9% SBFI: 10/60/600 OD: 11.7% TCD: -7.8%	EBFI: CBT/SP/NC/I/CO/NB SBFI: B/SP/NC/I/CO/NB Some home sessions with direct observation of parent-child interaction and feedback were included for both interventions.
^Sonuga-Barke et al., 2001 England	RCT; WLC	78 NFPP: 30 C&S: 28	36 mos.	Dx	Preschool equivalent of ADHD symptoms	New Forest Parenting Program (NFPP; Thompson Manual)+ I2: Parent Counseling and Support (C&S)+	Both NFP and C&S: 8/60/4780 OD: NP TCD: NP	NFPP: B/DI/C/I/HB/NB C&S: NB/DI/NC/I/HB/NB
~Dishion et al., 2008 United States	RCT; WLC	731 (364)	24-35 mos.	RS	Externalizing behavior problems	Family Check-Up+	3/NP/NP OD: 9.4% TCD: -0.2%	MI/SP/C/I/HB/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
~Posthumus et al.., 2012 Netherlands	Case control design; TAU	144 (72)	50.8 mos.	RS	Disruptive behavior problems/ conduct problems	Incredible Years Basic and Advanced Curriculum+	18/120/21 60 OD: 20.9% TCD: 34.5%	B/IP/NC/G/C/B
~Reid et al., 2013 Canada	RCT; TAU	178 (82)	38.4 mos.	PC	Externalizing behavior problems	Parenting Matters – Distance Based	6/NP/NP + 2 coach calls OD: 10.7% TCD: -9.6%	B/IP/NC/SA/HB*/NB *recruited from primary care
~Shaw et al., 2006 United States	RCT; TAU	120 (60)	24.1 mos.	RS	Externalizing behavior problems/ Aggression/ Disruptive behaviors	Family Check-Up+	3/NP/NP OD: 6.7% TCD: -3.3%	MI/SP/C/I/HB/NB
~Somech & Elizur, 2012 Israel	RCT; minimal interv. ctrl	209 (140)	48.57 mos.	O	Disruptive behavior problems/ conduct problems	Hitkashrut+	14/120/16 80 OD: 12.9% TCD: 6.7%	B/SP/C/G/C/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ Attrition[e]	Intervention Characteristics[f]
~Sourander et al., 2016 Finland	RCT; EC	464 (232)	48 mos.	RS	Disruptive behavior problems; antisocial behaviors	Strongest Families Smart Website (SFSW)+	11/45/495 OD: 14.7% TCD: -19.0%	B/IP/NC/SA/HB/B
~Van Zeijl et al., 2006 Netherlands	RCT; "fake" interv.	246; (120)	25.99 mos.	RS	Child externalizing behavior (opposition, aggression, overactivity)	Video-Feedback Method (VIPP-SD)+	6/90/540 OD: 2.1% TCD: -2.5%	NB/IP/C/I/HB/B

Note. [a] ^: in meta; ~:in systematic review [b] *Study design:* TAU: Treatment as Usual; WLC: Waitlist Control; EC: Educational Control; SC: Safety Control[c] *Eligibility* (RS = Had to be at or above clinical cutoff on parent self-reported measures or clinician/researcher screening questions/evaluation; PC = parent reported concern or self-referred (no rating scale cutoff or screening cutoff needed); Dx: have an age-corresponding, established diagnosis with active symptoms; NC = no pre-identified concern, diagnosis, or clinical cutoff needed, O= Other); [d] +:intervention is manualized. [e] *Number of Sessions/Minutes per session/Dosage:* First number is number of sessions, next is how many minutes per session, and last is dosage (number of sessions x minutes); OD = Overall Dropout Rate in %; TCD = Treatment/Control Differential in % - negative numbers mean greater attrition in control group, no sign otherwise means greater attrition in treatment group. Formulas for both are in appendix. [f] *Intervention Characteristics Grouping (in order): Theoretical Orientation* (B = Behavioral components; CBT = Cognitive Behavioral Components; MI = Motivational Interviewing; NB = Non-behavioral); *Type of prevention or intervention* (SP= selective Prevention; IP = Indicated Prevention; UI = Universal Intervention; DI = Direct Intervention); *Child Involvement* (C = Child involved; NC = Child not involved); *Delivery format* (I = Individual; G=Had group elements; SA = Self assisted or self-directed); *Treatment Setting* (HB = Home Based, IT = Internet, PC = Primary Care, CO = Community, U = University); *Booster* (B = Booster; NB = No Booster). Other Abbreviations: PI = Post-Intervention, NP = data not provided not able to calculate from data provided. Interv.=intervention. Comp. = comparison.

Table 3

Effect Sizes of Included Studies – Externalizing

Study [a]	Outcome Measures [b]	Hedges *g (SE)*	Time Outcomes Reported [c]	Long term effects reported?
^Abikoff et al., 2014 United States	ADHD-CPRS; NYPRS	HNC: -.68(.22) NFPP: -.49(.21)	PI	2 years for tx groups only
^Bradley et al., 2003 Canada	PCQ	-.34(.14)	PI	1 year f/up for tx group only (not in tables)
^Brassart et al., 2017 Belgium	CBCL	-.98(.33)	PI	4 mo for tx group only
^Brotman et al., 2003 United States	CBCL	-.68(.37)	PI	Only 6 families from tx group provided 6 mo data
^Dittman et al., 2016 New Zealand and Australia	ECBI	-.61(.22)	PI	6 mo f/up for tx grp only
^Hutchings et al., 2007 Wales/UK	ECBI, SDQ, Conners	-.48(.18)	PI	No
^Mackenzie and Hilgedick, 2000, United States	CBCL	CAPP: -.17(.36) Booklet: -.17(.36)	PI	1 month (all groups)
^Markie-Dadds and Sanders, 2006, Australia	ECBI	-.97(.32)	PI	6 mo for tx group only
^Morawska et al., 2011 Australia	ECBI	-.67(.29)	PI	6 mo for tx group only
^Morawska & Sanders, 2006 Australia	ECBI	SD-BFI: -.55(.26) TASD-BFI: -.64(.27)	PI	6 mos. for tx groups only

Study [a]	Outcome Measures [b]	Hedges *g (SE)*	Time Outcomes Reported [c]	Long term effects reported?
^Nicholson et al., 2002 United States	ECBI, PSC, BSQ	-.75(.40)	PI	1 month for tx grp only
^Nixon et al., 2001 United States	ECBI	.72(.35)	PI	6 mos. for tx and WLC
^Nixon et al., 2003 United States	ECBI, CBCL, ODD Sx, HSQ	STD: -.67(.34) ABB: -.47(.33)	PI	6, 12, and 24 mos. for tx groups only
^Perrin et al., 2014 United States	ECBI	R-IY: -.26(.17) NR-IY: -.31(.16)		across all groups for 6 and 12 mos.
^Sanders et al., 2000 Australia	ECBI	EBFI: -.81(.18) SBFI: -.68(.18)	PI	1 year f/up for tx groups only
^Sonuga-Barke et al., 2001 England	PSC	NFP:-1.13(.31) C&S:-.31(.29)	PI	23 weeks for both tx and control groups
~Dishion et al., 2008 United States	ECBI, CBCL	-.39(.08)	1 yr after intervention (age 3)	2 years after intervention (at age 4)
~Posthumus et al., 2012 Netherlands	ECBI	-.16(.20)	PI	1 and 2 years across all groups
~Reid et al., 2013 Canada	ECBI, CBCL	-.23(.16)	PI	across groups for 3 and 6 mos.; at 12 mos. for intervention group only
~Shaw et al., 2006 United States	CBCL	-.02(.19)	1 yr after intervention (age 3)	At age 4 (2 years PI)

Study [a]	Outcome Measures [b]	Hedges *g (SE)*	Time Outcomes Reported [c]	Long term effects reported?
~Somech and Elizur, 2012 Israel	ECBI, PCQ, ICU	-.43(.15)	PI	No
~Sourander et al., 2016 Finland	CBCL, ICU	-.26(.09)	6 mo PI	12 mo across all groups
~Van Zeijl et al., 2006 Netherlands	CBCL	-.09(.13)	8 mo. PI	No

Note. [a] ^: in meta; ~:in systematic review [b] *Rating scale abbreviations: ADHD-CPRS*; ADHD Connors Rating Scale-Revised; *BSQ:* Behavior Screening Questionnaire; *CBCL*: Child Behavior Checklist, Preschool Version (1.5-5 year olds); *EBCI*; Eyberg Child Behavior Inventory; *HSQ*: Home Situations Questionnaire, Modified (Severity of Problems Score); *ICU*: Inventory of Callous-Unemotional Traits (adapted for pre-school); *NYPRS*: New York Rating Scale (Opposition/Aggression Subscales); *ODD Sx*: Intensity of oppositional defiant disorder symptoms from Structured *DSM–IV* Interview; *PAC*: Parental Account of Childhood Symptoms; *PCQ:* Preschool Characteristics Questionnaire, Modified from Child Characteristics Questionnaire; *PSC:* Pediatric Symptom Checklist [c] *PI:* Post-intervention

Table 4

Meta-analytic results examining categorical moderators in response to externalizing interventions (TX1)

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Overall Model, w/o moderators	16	-.601	.071	-.740	-.463	-8.497	<.001	22.300	16.666	15	.339	72.204	1	<.001
Categorical Moderators														
Tx Setting								14.146	12.342	13	.500	3.532	2	.171
Community	10	-.524	.078	-.678	-.371	-6.696	<.001							
Home	3	-.364	.198	-.753	.024	-1.837	.066							
University	3	-.131	.189	-.503	.240	-.694	.488							
Delivery Format								0.000	10.496	13	.653	6.170	2	.046*
Individual	7	-.790	.109	-.1.004	-.575	-7.219	<.001							
Self-help	3	.167	.210	-.245	.579	.795	.427							
Group elements	6	.331	.134	.068	.594	2.468	0.014*							
Theoretical Orient.								9.861	12.633	13	.477	3.297	2	.192
Behavioral	12	-.533	.072	-.674	-.393	-7.446	<.001							
Cognitive-Beh.	3	-.244	.178	-.592	.104	-1.369	.171							
Non-behavioral	1	-.450	.345	-1.126	.227	-1.303	.193							
Interv./Prev.								14.602	11.830	12	.459	3.755	3	.289
Indicated prev.	7	-.518	.093	-.700	-.336	-5.564	<.001							
Selected prev.	4	-.138	.160	-.451	.175	-.864	.387							
Universal interv.	1	.349	.381	-.398	1.096	.915	.360							
Direct interv.	4	-.268	.180	-.621	.086	-1.484	.138							
Tx manualized?								25.135	16.646	14	.276	.000	1	.998
Yes	14	-.603	.075	-.750	-.457	-8.052	<.001							
No	2	-.001	.271	-5.320	.530	-.002	.998							

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Child involved?								15.633	14.099	14	.442	1.935	1	.164
Yes	5	-.771	.144	-1.054	-.489	-5.355	<.001							
No	11	.227	.163	-.093	.546	1.391	.164							
Sx Severity/ Child Eligibility								16.415	12.622	13	.477	3.091	2	.213
Parent concern/ self-referred	4	-.440	.124	-.682	-.197	-3.555	<.001							
Child had dx	5	-.332	.191	-.706	.042	-1.740	.082							
Sx/dx cutoff	7	-.174	.158	-.484	.136	-1.102	.270							
No dx/RS cutoff/ no parent concern	0	.	.		.	.	.							
Booster session?								15.593	13.628	14	.478	2.279	1	.131
Yes	4	-.415	.137	-.684	-.147	-3.035	.002							
No	12	-.238	.157	-.546	-.071	-1.509	.131							
Attrition in WWC acceptable range?								.000	7.284	9	.608	2.719	1	.099
Yes	5	-.646	.092	-.826	-.467	-7.050	<.001							
No	6	.214	.130	-.040	.467	1.649	.099							

Note. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Interv. = Intervention. Prev. = Prevention. RS = Rating Scale. WWC = What works Clearinghouse. Table 5 represents the 16 of the 23externalizing studies eligible to be analyzed by meta-analysis; see characteristics of studies table for further details.. If studies had a second intervention group, the second intervention group's data was used for the "TX2" analysis. For externalizing studies, if Hedges' g for the moderator is negative, the treatment group had less externalizing behaviors than the control group. Only 11 out 16 studies provided specific enough data to calculate attrition rate

Table 5

Meta-analytic results examining continuous moderators in response to externalizing interventions (TX1)

Continuous Moderators	*k*	*g*	*SE*	*95%CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*P*	*QO*	*df*	*p*
Child age, mos.	16	.003	.013	-.023	.029	-1.300	.193	27.712	16.421	13	.227	.058	1	.810
Number of Tx Sessions	15	-.004	.007	-.018	.010	-.539	.590	21.074	14.578	13	.334	.291	1	.590
Session Duration	15	.004	.002	.000	.007	2.182	.029*	.000	10.327	13	.667	4.761	1	.029*
Dosage in minutes	15	.000	.000	.000	.000	.092	.927	23.543	15.075	13	.303	.008	1	.927
Overall Dropout Rate	11	-.511	1.195	-2.853	1.832	-.427	.669	23.972	9.826	9	.365	.183	1	.669
Control/ Treatment Differential Dropout Rate	11	.781	.592	-.379	1.940	1.320	.187	3.921	8.064	9	.528	1.742	1	.187

Note. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Tx = Treatment. Table 5 represents the 16 of the 23 externalizing studies eligible to be analyzed by meta-analysis; see characteristics of studies table for further details. If studies had a second intervention group, the second intervention group's data was used for the "TX2" analysis. For externalizing studies, if Hedges' g for the moderator is negative, the treatment group had less externalizing behaviors than the control group.

Table 6

Meta-analytic results examining categorical moderators in response to externalizing interventions (TX2)

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*Df*	*p*	*QO*	*Df*	*p*
Overall Model w/o moderators	16	-.510	.058	-.625	-.395	-8.721	<.001	.000	11.121	15	.744	76.060	1	<.001
Categorical Moderators														
Tx Setting								.000	10.184	13	.679	.937	2	.626
Community	8	-.483	.073	-.626	-.340	-6.611	< .001							
Home	5	-.022	.143	-.302	.259	-.150	.880							
University	3	-.171	.177	-.518	.176	-.967	.334							
Delivery Format								.000	10.529	13	.650	.592	2	.744
Individual	5	-.557	.110	-.772	-.343	-5.087	< .001							
Self-help	3	-.027	.207	-.432	.378	-.130	.897							
Group	8	.083	.133	-.177	.344	.626	.532							
Theoretical Orientation								.000	10.291	13	.670	.830	2	.660
Behavioral	12	-.492	.062	-.614	-.370	-7.898	< .001							
Cognitive-Beh.	2	-.220	.280	-.770	.329	-.787	.431							
Non-behavioral	2	-.115	.227	-.560	.330	-.506	.613							
Int./prev.								.000	8.641	12	.733	2.48	3	.479
Indicated prev.	8	-.489	.073	-.632	-.346	-6.700	< .001							
Selected prev.	3	-.198	.165	-.522	.125	-1.202	.229							
Universal int.	1	.320	.363	-.391	1.031	.882	.378							
Direct int.	4	.007	.158	-.302	.317	.047	.962							
Tx manualized?								.000	10.943	14	.690	.178	1	.673
Yes	14	-.504	.060	-.622	-.386	-8.358	< .001							
No	2	-.105	.248	-.592	.382	-.422	.673							

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*Df*	*p*	*QO*	*Df*	*p*
Sx Severity/ Child Eligibility								.000	9.835	13	.707	1.286	2	.526
Parent concern/ Self-referred	4	-.419	.104	-.623	-.214	-4.012	< .001							
Child had dx	5	-.088	.167	-.416	.241	-.523	.601							
Sx or dx cutoff	7	-.152	.134	-.414	.111	-1.134	.257							
No dx/RS cutoff/ no parent concern	0	.	.		.	.	.							
Booster session?								.000	9.393	14	.805	1.728	1	.189
Yes	4	-.378	.116	-.606	-.150	-3.248	.001							
No	12	-.177	.135	-.441	.087	-1.314	.189							
Attrition in WWC acceptable range?								.000	6.141	9	.726	.750	1	.387
Yes	2	.128	.148	-.162	.418	.866	.387							
No	8	-.531	.073	-.673	.388	-7.295	<.001*							

Note. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Tx. = Treatment. Int. = Intervention. Prev. = Prevention. RS = rating scale. Sx = Symptom. Dx = Diagnosis. WWC = What Works Clearinghouse. Table 6 represents the 16 of the 23 externalizing studies eligible to be analyzed by meta-analysis; see characteristics of studies table for further details. If studies had a second intervention group, the second intervention group's data was used for the "TX2" analysis. For externalizing studies, if Hedges' g for the moderator is negative, the treatment group had less externalizing behaviors than the control group. Only 11 out 16 studies provided specific enough data to calculate attrition rates.

Table 7

Meta-analytic results examining continuous moderators in response to externalizing interventions (TX2)

Continuous Moderators	*k*	*g*	*SE*	*95%CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Child age, mos.	16	-.005	.011	-.026	.016	-.481	.630	.000	10.531	13	.650	.232	1	.630
Number of Treatment Sessions	15	-.004	.000	-.017	.009	-.580	.562	.000	9.823	12	.631	.336	1	.562
Session Duration	13	.003	.002	-.002	.007	1.185	.236	.000	6.671	10	.756	1.403	1	.236
Dosage in minutes	13	.000	.000	.000	.000	-.181	.856	.000	8.052	11	.709	.033	1	.856
Overall Dropout Rate	11	.220	.697	-1.147	1.587	.316	.752	.000	.100	1	.659	1.100	1	.752
Control/Treatment Differential Dropout Rate	11	-.033	.555	-1.121	1.056	-.059	.953	.000	6.888	9	.649	.003	1	.953

Note. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Table 7 represents the 16 of the 23 externalizing studies eligible to be analyzed by meta-analysis; see characteristics of studies table for further details. If studies had a second intervention group, the second intervention group's data was used for the "TX2" analysis. For externalizing studies, if Hedges' g for the moderator is negative, the treatment group had less externalizing behaviors than the control group.

Figure 2

Forest Plot – Externalizing Interventions Included in First Meta-Analysis (TX1)

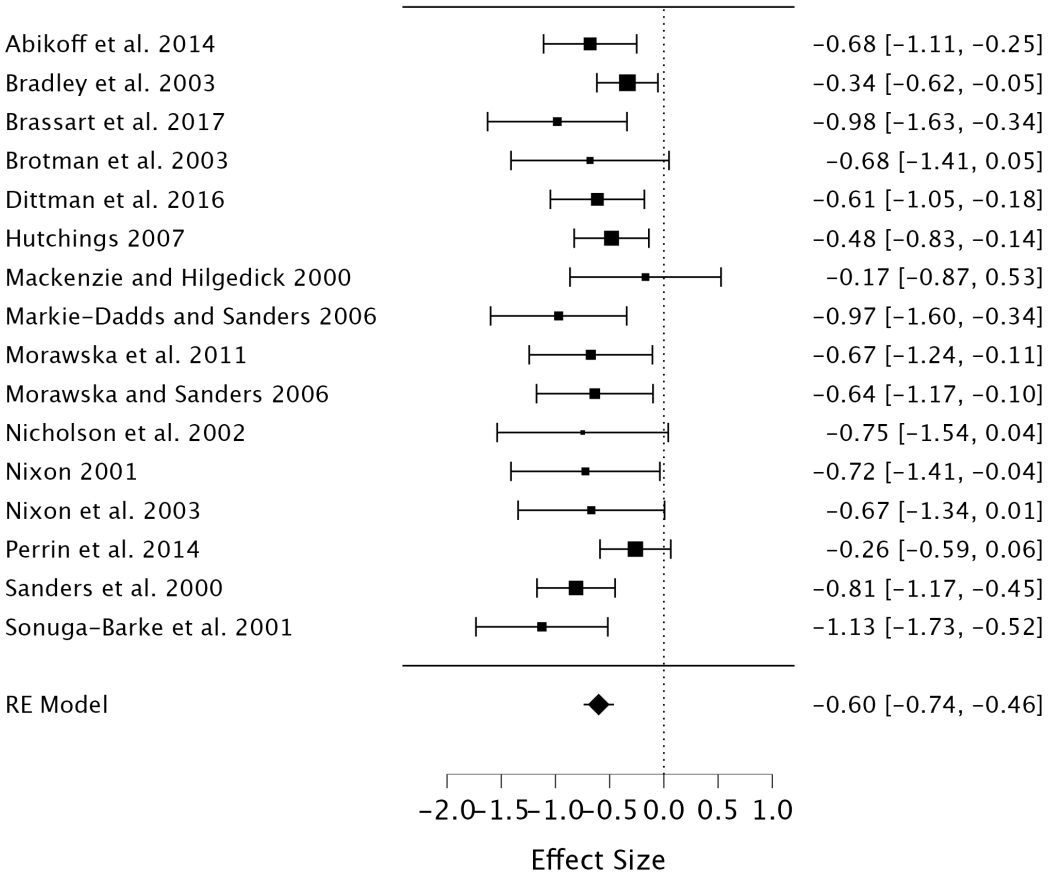

Figure 3

Funnel Plot – Externalizing Interventions Included in First Meta-Analysis (TX1)

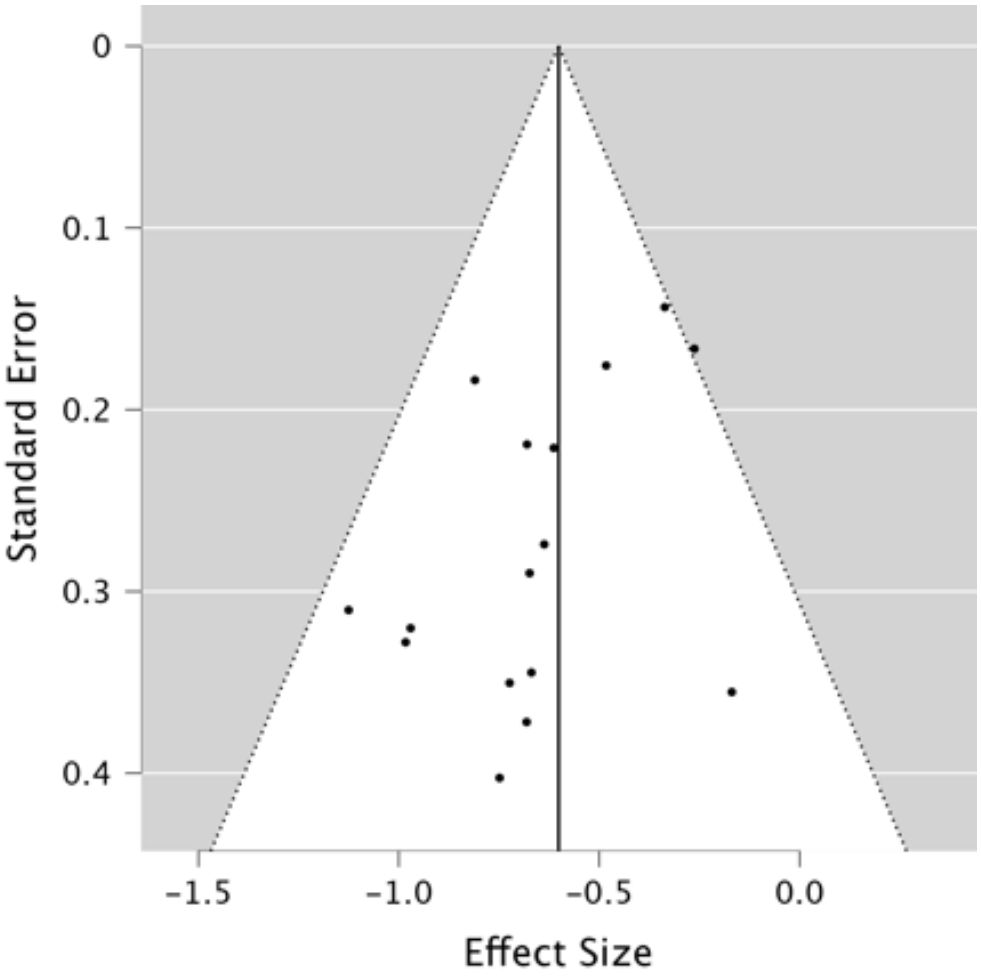

Figure 4

Forest Plot – Externalizing Interventions Included in Second Meta-Analysis (TX2)

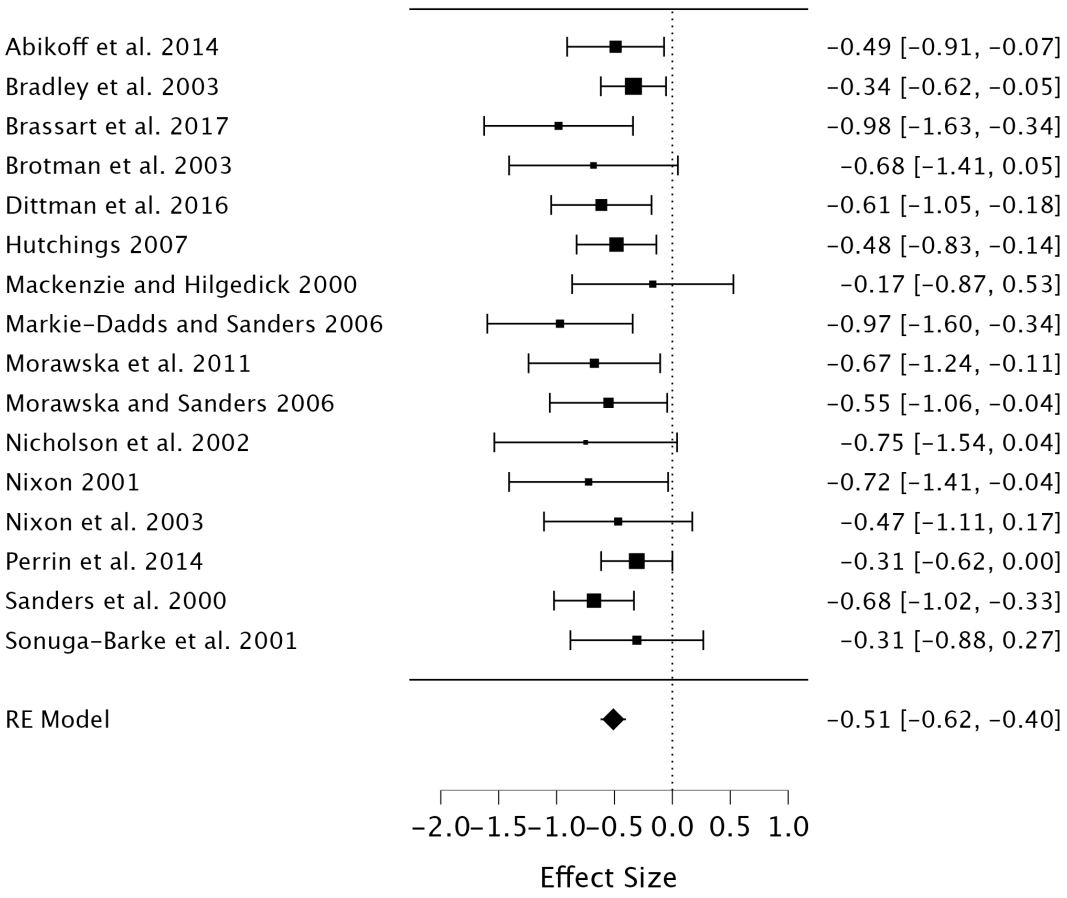

Figure 5

Funnel Plot – Externalizing Interventions Included in Second Meta-Analysis (TX2)

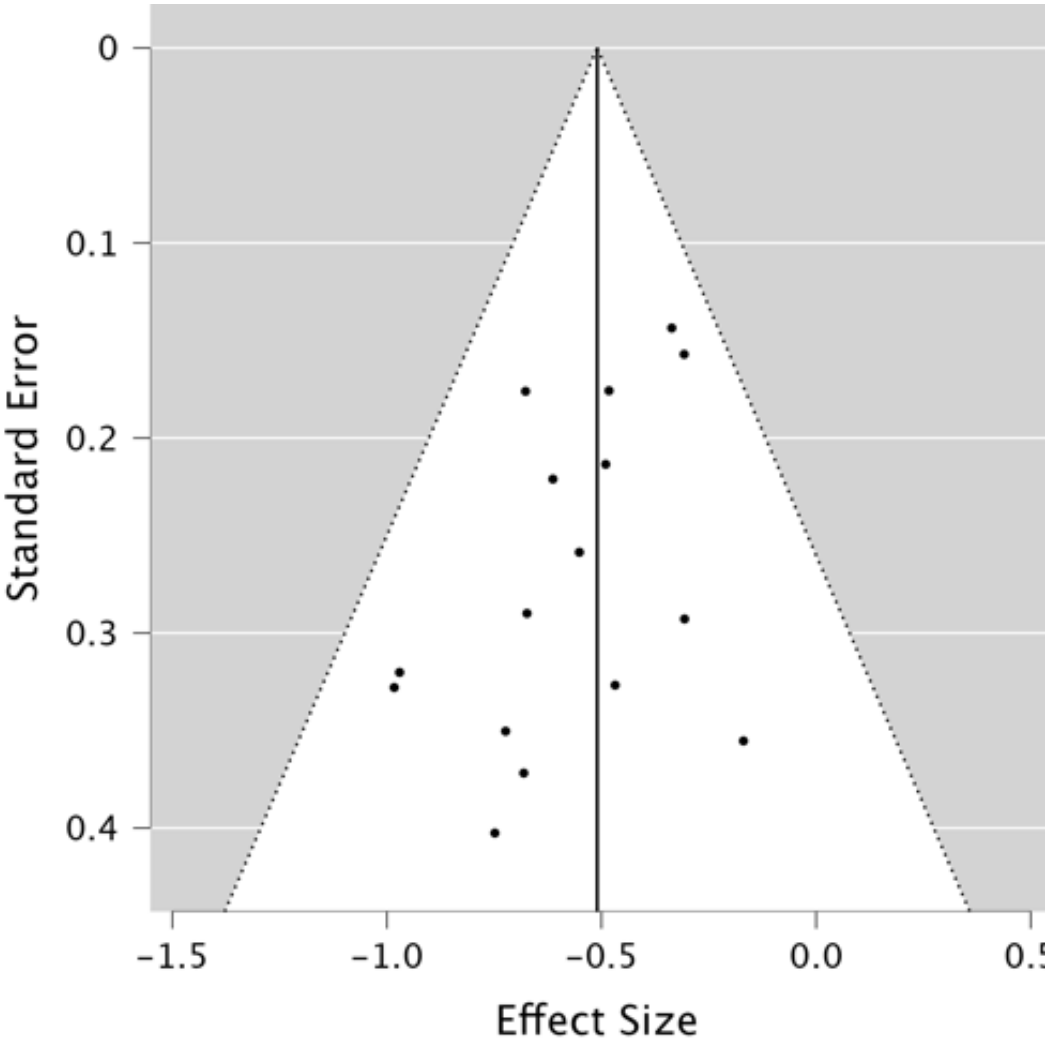

Table 8

Descriptive statistics – All internalizing studies

	k	*M*	*SD*	*Range*
Parent age (yrs.)	2	36.608	0.576	36.200 - 37.015
Child age (mos.)	3	48.697	3.344	46.450 - 52.540
Percent female parent	1	0.587	.	.587 - .587
Percent white parent	2	0.611	0.077	.556 - .665
Percent $\geq$ college education	1	0.512	.	.512 - .512
Percent female child	2	0.503	0.110	.425 - .580
Overall attrition	3	0.174	0.187	0 - .372
Treatment vs. control differential attrition	3	-0.048	0.108	-0.172 - .027

Table 9

Characteristics of Included Studies – Internalizing

Study	Study Design/ Control Type [a]	Total *n*/ (Tx *n*)	Child's Mean age	Eligibility[b]	Target Behavior	Intervention Name[c]	Tx length/ attrition [d]	Intervention Characteristics [e]
Chronis-Tuscano et al., 2015 United States	RCT; WLC	40 (18)	44 mos.	RS	Behavioral inhibition and anxiety symptoms	The Turtle Program+(adapted from PCIT for Parents) and Social Skills Facilitated Play for Kids	8/90/720 OD = 15% TCD = -17%	B/IP/C/G/C/NB
Kennedy et al., 2009 Australia	RCT; WLC	71 (35)	46.8 mos.	RS	Behavioral inhibition and anxiety symptoms	Modification of the Cool Kids Program+	8/90/720 NP	CBT/SP/NC/G/U/B
Rapee et al., 2005 Australia	RCT; TAU	146 (73)	46.8 mos.	RS	Behavioral inhibition and anxiety symptoms	"Parent early intervention program" (Cool Kids Program+)	6/90/540 OD = 37.2% TCD = 2.7%	CBT/IP/NC/G/U/NB

Note. [a] *Study design:* TAU: Treatment as Usual; WLC: Waitlist Control; EC: Educational Control; SC: Safety Control [b] *Eligibility*: RS = Had to be at or above clinical cutoff on parent self-reported measures or clinician/researcher screening questions/evaluation [c] +indicates manualized intervention[d] *Number of Sessions/Minutes per session/Dosage:* First number is number of sessions, next is how many minutes per session, and last is dosage (number of sessions x minutes); OD = Overall Dropout Rate in %; TCD = Treatment/Control Differential in % - negative numbers mean greater attrition in control group, no sign otherwise means greater attrition in treatment group. Formulas for both are in appendix. [e] *Intervention Characteristics Grouping (in order): Theoretical Orientation* (B = Behavioral components; CBT = Cognitive Behavioral Components; *Type of prevention or intervention* (SP= selective Prevention; IP = Indicated Prevention; *Child Involvement* (C = Child involved; NC = Child not involved); *Delivery format* (G=Had group elements); *Treatment Setting* (C = Community, U = University); *Booster* (B = Booster; NB = No Booster). Other Abbreviations: PI = Post-Intervention, NP = data not provided not able to calculate from data provided. Interv.=intervention

Table 10

Effect Sizes of Included Studies – Internalizing

Study	Outcome Measures	Hedges *g (SE)* and time reported	Long term effects reported?
Chronis-Tuscano et al., 2015, United States	Parent Report of PAS, BIQ, and CBCL Internalizing	-3.470 (.728) post-intervention	No
Kennedy et al., 2009 Australia	Maternal and paternal reports of BIQ	-.628 (.258) 6 mo after intervention	No
Rapee et al., 2005 Australia	Maternal report - PAS Maternal and paternal reports on the Temperament Assessment Battery for Children, Revised, Inhibition Subscale	-.058 (.182) 12 mo after intervention	24 mos. and 36 mos. (monitored preschool children for 11 years Rapee 2013)

Note. BIQ = Behavioral Inhibition Questionnaire; PAS-R = Preschool Anxiety Scale, Revised; CBCL= Child Behavior Checklist, CALIS = Child Anxiety Life Interference Scale

Table 11

Descriptive statistics – All sleep studies

	k	*M*	*SD*	*Range*
Parent age (yrs.)	5	31.974	2.538	28.700 -35.375
Child age (mos.)	10	17.931	14.692	0 - 48.000
Percent female parent	7	.999	.004	.990 - 1.000
Percent white parent	5	.859	.105	.681 - .940
Percent $\geq$ college education	8	.793	.189	.400 - .943
Percent female child	10	.499	.034	.443 - .544
Overall attrition	8	.1174	.0715	.0175 - .2290
Treatment vs. control differential attrition	8	.0152	.1476	-.2500 - .2993

Table 12

Characteristics of Included Studies – Sleep

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ attrition[e]	Intervention Characteristics[f]
^Mindell et al., 2009 (Infant Study) United States	RCT; TAU	206 (134)	18 mos.	PC	sleep onset latency, reducing disruptive bedtime behavior	3-step bedtime routine w/o other behavioral intervention (bath, massage, quiet activities)	NP b/c SA, though occurred over course of 3 weeks OD: NP TCD: NP	CSP only: E/IP/NC/SA/IT/NB CSP+BR: B/IP/NC/SA/IT/NB
^Mindell et al., 2009 (Toddler Study) United States	RCT; TAU	199 (133)	36 mos.	PC	sleep onset latency, reducing disruptive bedtime behavior	3-step bedtime routine without other behavioral intervention (bath, massage, quiet activities)	NP b/c SA, though occurred over course of 3 weeks OD: NP TCD: NP	B/IP/NC/SA/IT/NB
^Mindell et al., 2011 United States	RCT; WLC	272 CSP= 96 CSP +BR= 84	30.95 mos. (range is 6-36 mos.)	PC	Sleep disturbance	Internet Based Intervention – Customized Sleep Profile (CSP) Internet Based Intervention – Customized Sleep Profile + Bedtime Routine(CSP&BR)	NP b/c SA, though occurred over course of 3 weeks OD: 1.8% TCD: -1.1%	CSP only: E/IP/NC/SA/IT/NB CSP+BR: B/IP/NC/SA/IT/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ attrition[e]	Intervention Characteristics[f]
^Reid et al., 1999 United States	RCT; WLC	41 S =12 GE = 13	48 mos.	RS	Bedtime problems	Scripted treatment manuals for: +Standard Ignoring Group (SI) +Graduated Extinction Group (GE)	2 telephone sessions for both grps; calls only if issues arose over course of 2 weeks SI: OD= 12.5% TCD= -25.0% GE: OD= 12.1% TCD= -23.5%	For both interventions: B/IP/NC/I/NC/HB/NB
^Schlarb et al., 2017 Germany	RCT; WLC	199 MK+TS =72 MK w/o TS = 58	19.71 mos.	PC	Sleep disturbances (sleep latency, duration of night awake, general sleep problems	+Mini-Kiss with telephone support from psychologists (MK&TS) +Mini-Kiss without telephone support (MK w/o TS)	MK+TS 6/10/60 OD = 7.1% TCD= 0.3% MK w/o TS Tx length NP due to SA OD= 11.8% TCD= 10.0%	MK+TS: CBT/IP/NC/ID/IT/NB MK w/o TS: CBT/IP/NC/SA/IT/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ attrition[e]	Intervention Characteristics[f]
^Stevens et al., 2019 United States	RCT; WLC	239 DVD = 80 CSP = 82	17.7 mos.	PC	Falling and staying asleep	DVD group: Extinction and parent education via Sleep Easy Solution DVD CSP group: Sleep education via Customized Sleep Profile from Johnson and Johnson Baby Sleep Website	NP due to SA DVD: OD = 18.5% TCD = 2.0% CSP: OD = 25.8% TCD = -12.2%	DVD group: B/IP/NC/SA/HB/NB CSP group: E/IP/NC/SA/IT/NB
^St. James-Roberts et al., 2001 United Kingdom	RCT; TAU	610 BP = 205 EP = 202	Recruited after birth	NC	general bedtime problems, reduce night waking	Behavioral Program (BP) leaflet and researcher discussion graduated extinction and bedtime routine Educational program (EP)	BP: 1/NP/NP OD = NP TCD = NP EP: NP due to SA	BP: B/UI/NC/I/HB/B EP: E/UI/NC/SA/HB/NB
^Stremler et al., 2013 Canada	RCT; TAU	246 (123)	Recruited after birth	NC	reduce night waking over time	Educational program - "infant sleep structure, sleep promotion, & differentiating b/t night/ day"	3/60/180 OD= 16.7% TCD= 0.8%	E/UI/NC/I/HB/NB

Study [a]	Study Design[b]	Total *n* (Tx *n*)	Child's mean age	Eligibility[c]	Target Behavior	Intervention[d]	Tx length/ attrition[e]	Intervention Characteristics[f]
~Adachi et al., 2009 Japan	PC; EC	196 (99)	4 mos.	NC	Promoting favorable sleep patterns; preventing sleep disturbance	"Baby, Sleep Well at Night" informational packet + 10 minute group presentation	1/10/10 NP	E/UI/NC/G/PC/NB
~Eckerberg, 2002 Sweden	QE; EC	67 AS = 39 WI = 28	9.8 mos.	PC	reduce night waking over time	AS: Two step variation graduated extinction w/ advice and support WI: Two step variation graduated extinction – written information only	AS: 2 sessions totaling 90 min; add'l 120 min in support calls WI: written information only NP	B/IP/NC/I/HB/NB
~Hall et al., 2015 Canada	RCT; SC	235 (117)	6.75 mo (6-8 month range)	RS	reduce night waking over time	Two-Hour Nurse led sleep education group (setting limits, routines, normal infant sleep, etc)	1/120/120 + 4 support calls DO: 8.5% TCD: -1.63%	E/IP/NC/G/PC/NB
~Paul et al., 2016 United States	RCT; SC	279 (140)	Random-ized after birth	NC	"improve sleep behaviors and duration"	Insight Responsive Parenting Interv. education on developmentally appropriate sleep	4/NP/NP OD: 6.1% TCD: 3.6%	E/UI/NC/I/HB/NB

Note. [a] ^: in meta; ~:in systematic review [b] *Study design:* TAU: Treatment as Usual; WLC: Waitlist Control; EC: Educational Control; SC: Safety Control QE: Quasi-Experimental [c] *Eligibility* (RS = Had to be at or above clinical cutoff on parent self-reported measures or clinician/researcher screening questions/evaluation; PC = parent reported concern or self-referred (no rating scale cutoff or screening cutoff needed); NC = no pre-identified concern, diagnosis, or clinical cutoff needed);[d] +:intervention is manualized. [e] *Number of Sessions/Minutes per session/Dosage:* First number is number of sessions, next is how many minutes per session, and last is dosage (number of sessions x minutes); OD = Overall Dropout Rate in %; TCD = Treatment/Control Differential in % - negative numbers mean greater attrition in control group, no sign otherwise means greater attrition in treatment group. Formulas for both are in appendix. [f]*Intervention Characteristics Grouping (in order): Theoretical Orientation* (B = Behavioral components; E=Psychoeducation/Educational; *Type of prevention or intervention* (SP= selective Prevention; IP = Indicated Prevention; UI = Universal Intervention; DI = Direct Intervention); *Child Involvement* (C = Child involved; NC = Child not involved); *Delivery format* (I = Individual; G=Had group elements; SA = Self assisted or self-directed); ***Treatment Setting*** (HB = Home Based, IT = Internet, PC = Primary Care, CO = Community, U = University); *Booster* (B = Booster; NB = No Booster). Other Abbreviations: PI = Post-Intervention, NP = data not provided not able to calculate from data provided. Interv.=intervention

Table 13

Effect Sizes of Included Studies – Sleep

Study [a]	**Outcome Measures** [b]	**Hedges *g (SE)***	**Time Outcomes Reported** [c]	**Long term effects reported?**
^Mindell et al., 2009 (Infant Study) United States	BISQ; sleep diary	.45(.15)	PI	No
^Mindell et al., 2009 (Toddler Study) United States	BISQ; sleep diary	.35(.23)	PI	No
^Mindell et al., 2011 United States	BISQ	.46(.15)	PI	Yes; 1 year PI across all groups
^+Reid et al., 1999 United States	GB, GN, CBCL 1.5-5 Sleep Subscale	GE: 3.34 (.66) SI: 2.38 (.16)	PI	Yes, 2 months PI across all groups
^+Schlarb et al., 2017 Germany	sleep diary; MKO-AQ CBCL 1.5-5 Sleep Subscale	MK+TS: 0.69 (.18) MK w/o TS: 0.52 (.03)	PI	No
^Stevens et al., 2019 United States	BISQ	DVD: .25 (.16) CSP: -.04 (.03)	1 month PI	No
^St. James-Roberts et al.. 2001 United Kingdom	Sleep diary	BP: .01(0.1) EP: -.03(0.1)	PI	9 months for tx groups only
^Stremler et al., 2013 Canada	Actigraphy	.17(.14)	PI	No

Study [a]	Outcome Measures [b]	Hedges *g (SE)*	Time Outcomes Reported [c]	Long term effects reported?
~Adachi et al., 2009 Japan	Questionnaire created for study (sleep related problems)	.07(.73)	3 months PI	No
~Eckerberg 2002, Sweden	Sleep diary	.13(.25)	PI	Yes; 1 and 3 mos. PI
~Hall et al., 2015 Canada	Sleep diary	-.002(.21)	6 weeks PI	No
~Paul et al., 2016 United States	BISQ	.01(.28)	PI	No

Note. [a] ^: in meta; ~:in systematic review [b] *Rating Scale Abbreviations: BISQ:* Brief Infant Sleep Questionnaire; *GB:* Good bedtimes; child settled alone <10 minutes; *GN:* Good nighttimes5- child slept through night without waking parents; *CBCL:* Child Behavior Checklist; MKO-AQ: Mini Kiss Online Anamnestic Questionnaire. *Other abbreviations*: *GE:* Graduated Extinction; *SI:* Standard Ignoring; *MK:* Mini-Kiss Intervention; *TS:* telephone support; *BP:* behavior program; [c]PI: post-intervention.

Table 14

Meta-analytic results examining categorical moderators in response to sleep interventions (TX1)

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Overall Model, without moderators	8	.415	.145	.129	.700	2.849	.004*	85.151	35.023	7	<.001	8.118	1	004*
Categorical Moderators														
Tx Setting								88.32	22.557	6	<.001	.106	1	.744
Home, other	4	.377	.244	-.102	.855	1.544	.123							
Home, internet	4	.109	.335	-.548	.767	.326	.744							
Delivery Format								90.963	31.140	6	<.001	.183	1	.668
Self-guided	4	.376	.262	-.136	.889	1.439	.150							
Individual	4	.163	.381	-.584	.910	.428	.668							
Theoretical Orientation								86.408	31.091	6	<.001	1.103	1	.294
Behavioral	7	.488	.167	.162	.815	2.930	.003*							
Psychoeducation	1	-.472	.450	-1.354	.409	-1.050	.294							
Interv./prev.								.009	16.996	6	.009	18.025	1	<.001*
Indicated prev.	6	.463	.070	.326	.601	6.619	<.001*							
Universal interv.	2	-.453	.107	-.662	-.244	-4.246	<.001*							
Treatment manualized?								70.293	21.084	6	.002	7.772	1	.005*
No	6	.244	.111	.027	.461	2.204	.028*							
Yes	2	.787	.282	.234	1.340	2.788	.005*							
Child involved?~														
Yes	0	.	.		.	.	.	.	.	.				
No	8	.	.		.	.	.	.	.	.				

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Symptom Severity/Child Eligibility								.000	3.882	5	.556	31.141	2	<.001*
Sx or dx cutoff needed	1	2.376	.533	1.332	3.420	4.460	<.001*							
No dx/RS cutoff/no parent concern	5	-2.365	.539	-3.421	-1.309	-4.390	<.001*							
Parents reported concern only	2	-1.946	.537	-2.999	-.893	-3.621	<.001*							
Booster session?								83.257	25.382	6	p<.001	1.312	1	.252
No	7	.487	.160	.173	.801	3.044	.002*							
Yes	1	-.479	.418	-1.298	.340	-1.146	.252							
Attrition within WWC acceptable range?								70.875	10.394	3	.015*	11.341	1	<.001*
Yes	4	-2.034	.604	-3.217	-.850	-3.368	<.001*							
No	1	2.376	.586	1.227	3.515	4.053	<.001*							

Note. ~:no studies involved children directly. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Interv = Intervention. Prev. = Prevention. RS = Rating Scale. WWC = What Works Clearinghouse. Table 14 reflects the 8 studies that were included in sleep meta-analytic analysis only. The other 4 studies in systematic review are described in the study characteristics table. For sleep specifically, if Hedges' g or beta value for the moderator is positive, treatment group had better/improved sleep compared to the control. If negative, control group had better sleep. Only five of the eight studies in this meta-analysis provided specific data to calculate attrition rates.

Table 15

Meta-analytic results examining continuous moderators in response to sleep interventions (TX1)

	k	*g*	*SE*	*95%CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Child age, mos.	8	.028	.010	.009	.047	2.831	.005*	71.74	19.438	6	.003	8.015	1	.005*
Number of Tx Sessions	4	-.091	.307	-.693	.512	-.295	.768	97.90	22.706	2	<.001	.087	1	.768
Session Duration	2	.	.		.	.	.	.	.					.
Dosage in minutes	2	.	.		.	.	.	.	.					.
Overall Dropout Rate	5	-1.632	6.531	-14.431	11.168	-0.250	.8043	95.696	19.721	3	<.001	.062	1	.803
Control/ Tx Differential Dropout Rate	5	-8.069	2.260	-12.499	-3.639	-3.570	.001*	63.529	8.318	3	.040	12.744	1	.001

Note. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Tx = Treatment Table 15 reflects the 8 studies that were included in sleep meta-analytic analysis only. The other 4 studies in systematic review are described in the study characteristics table. For sleep specifically, if Hedges' g or beta value for the moderator is positive, treatment group had better/improved sleep compared to the control. If negative, control group had better sleep. Only five of the eight studies in this meta-analysis provided specific data to calculate attrition rates.

Table 16

Meta-analytic results examining categorical moderators in response to sleep interventions (TX2)

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Overall Model, without moderators	8	.460	.250	-.032	.948	1.832	.067	95.050	41.174	7	<.001	3.356	1	.067
Categorical Moderators														
Tx Setting	8							96.493	37.938	6	<.001	.481	1	.488
Home, other	4	.717	.441	-.147	1.581	1.626	.104							
Home, internet	4	-.423	.609	-1.617	.772	-.694	.488							
Delivery Format	8							95.853	41.029	6	<.001	2.023	1	.155
Self -guided	6	.275	.312	-.337	.886	.881	.378							
Individual	2	.968	.681	-.366	2.302	1.422	.155							
Theoretical Orientation	8							94.658	32.142	6	<.001	2.630	1	.105
Behavioral	3	1.009	.421	.185	1.834	2.398	.016*							
Psychoedu.	5	-.837	.516	-1.849	.175	-1.622	.105							
Interv./Prev.	8							95.206	28.182	6	<.001	1.190	1	.275
Indicated prev.	6	.645	.310	.037	1.253	2.080	.038*							
Universal interv.	2	-.653	.599	-1.827	.520	-1.091	.275							
Tx manualized?	8							92.181	31.690	6	<.001	5.371	1	.020*
No	6	.191	.218	-.235	.618	.879	.379							
Yes	2	1.171	.505	.181	2.162	2.318	.020							

	k	*g*	*SE*	*95% CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Child involved?~														
Yes	0	.	.	.		.	.	.				.		
No	8	.	.	.		.	.	.				.		
Sx Severity/Child Eligibility	8							27.981	7.408	2	<.001*	29.103	2	<.001*
Sx or dx cutoff	1	3.337	.661	2.041	4.633	5.046	<.001*							
No dx/RS cutoff no parent concern	5	-3.350	.670	-4.662	-2.037	-5.003	<.001*							
Parent concern only	2	-3.002	.666	-4.308	-1.696	-4.505	<.001*							
Booster session?^	8													
No	8	.	.		.	.	.	.				.		
Yes	0	.	.		.	.	.	.				.		
Attrition within WWC acceptable range?								98.018	31.183	3	<.001*	.556	1	.452
Yes	2	-.922	1.226	-3.324	1.480	-.752	.452							
No	3	1.136	.791	-.414	2.686	1.436	.151							

Note. ~:no studies involved children directly. ^:no studies in TX2 had booster sessions. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Interv. = Intervention. Psychoedu. = Psychoeducation. WWC = What Works Clearinghouse. Table 16 reflects the 8 studies that were included in sleep meta-analytic analysis only. The other 4 studies in systematic review are described in the study characteristics table. For sleep specifically, if Hedges' g or beta value for the moderator is positive, treatment group had better/improved sleep compared to the control. If negative, control group had better sleep. Only five of the eight studies in this meta-analysis provided specific data to calculate attrition rates.

Table 17

Meta-analytic results examining continuous moderators in response to sleep interventions (TX2)

Continuous Moderators	*k*	*g*	*SE*	*95%CI*		*z*	*p*	I^2	*QB*			*QO*		
				LL	*UL*				*QB*	*df*	*p*	*QO*	*df*	*p*
Child age, mos.	8	.066	.027	.012	.120	2.410	.016*	92.495	24.372	4	<.001	5.809	1	.016*
Number of Tx Sessions	4	-.219	.429	-1.060	.622	-.510	.610	98.875	27.558	2	<.001	.260	1	.610
Session Duration	2	.	.	.		.	.	.						.
Dosage in minutes	2	.	.	.		.	.	.						.
Overall Dropout Rate	5	-4.114	8.450	-20.675	12.446	-.487	.626	98.167	24.784	3	<.001	.237	1	.626
Control/ Treatment Differential Dropout Rate	5	-11.073	5.030	-20.933	-1.214	-2.201	.028*	94.981	27.201	3	<.001	4.846	1	.028*

Note. CI = Confidence Interval; LL = lower limit; UL = upper limit. QB= Test of Residual Heterogeneity. QO = Omnibus Test of Model Coefficients. Table 17 reflects the 8 studies that were included in sleep meta-analytic analysis only. The other 4 studies in systematic review are described in the study characteristics table. For sleep specifically, if Hedges' g or beta value for the moderator is positive, treatment group had better/improved sleep compared to the control. If negative, control group had better sleep.

Figure 6

Forest Plot of Sleep Interventions Included in First Meta-Analysis (TX1)

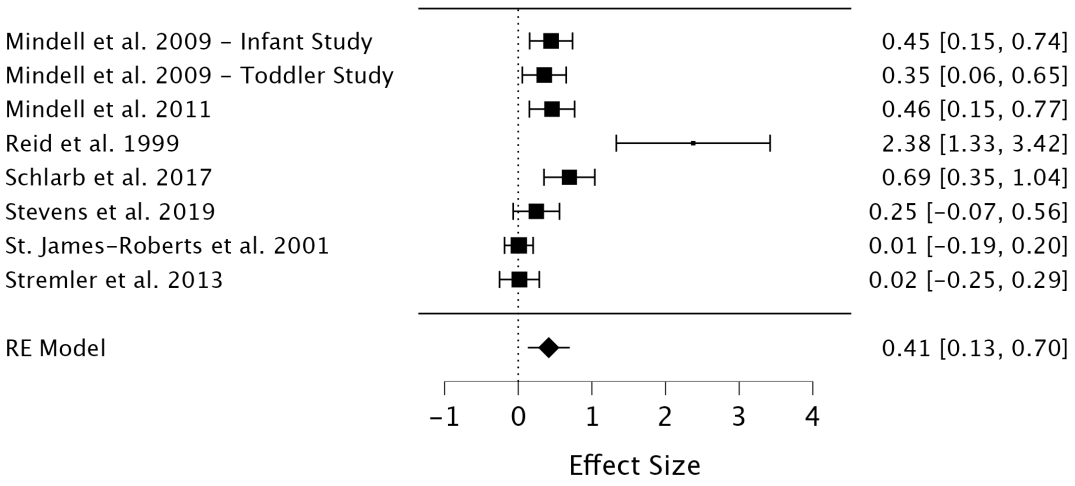

Figure 7

Funnel Plot of Sleep Interventions Included in First Meta-Analysis (TX1)

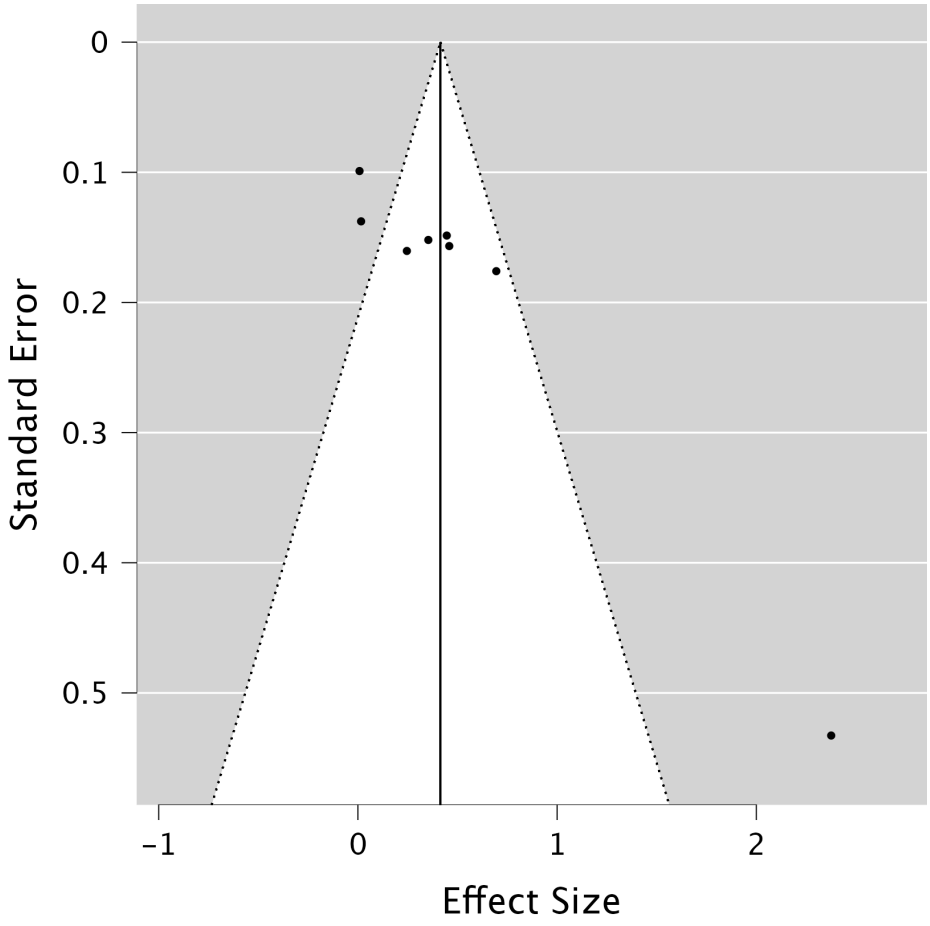

Figure 8

Forest Plot of Sleep Interventions Included in Second Meta-Analysis (TX2)

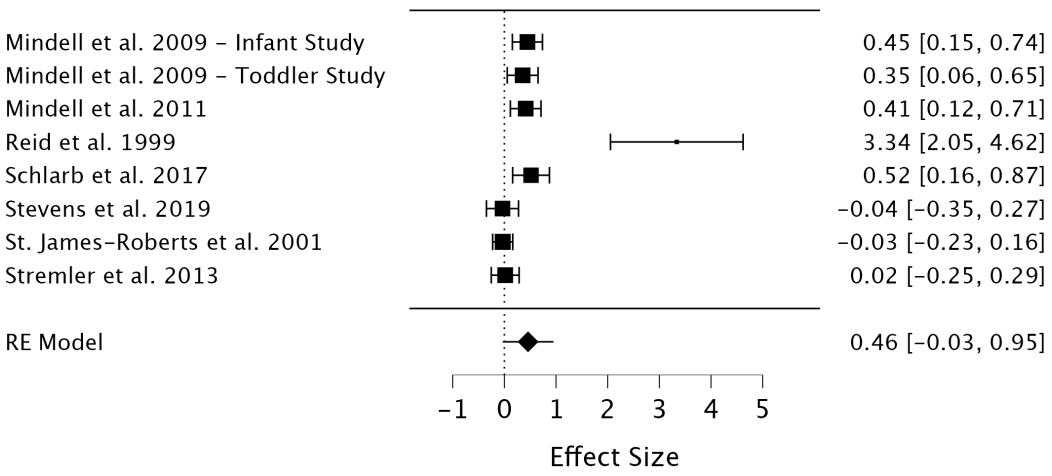

Figure 9

Funnel Plot of Sleep Interventions Included in Second Meta-Analysis (TX2)

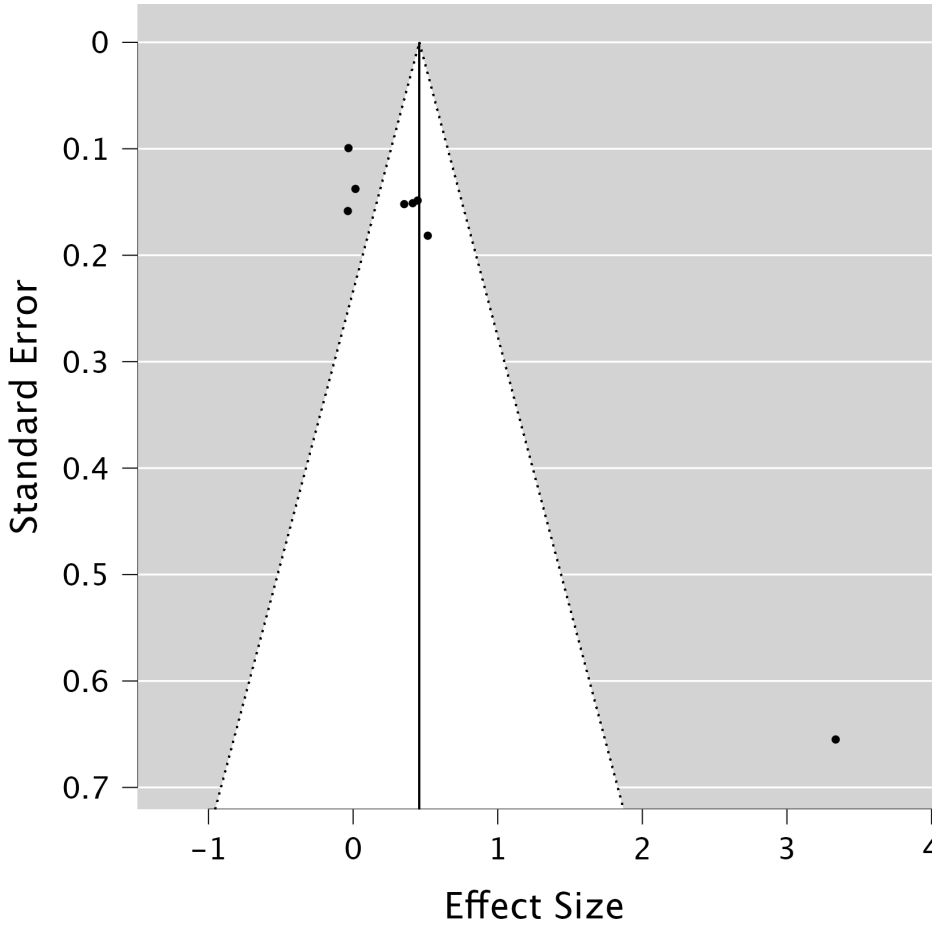

Table 18

Descriptive statistics – All feeding studies

	k	*M*	*SD*	Range
Parent age (yrs.)	3	32.114	7.524	23.565-37.728
Child age (mos.)	3	30.497	15.527	13.87 - 44.62
Percent female parent	2	.480	.537	.100 - .860
Percent white parent	2	.698	.361	0.443 -.953
Percent $\geq$ college education	3	.408	.315	.044 - .600
Percent female child	2	.513	.014	.503 - .523
Overall attrition	3	.111	.060	.044 -.160
Treatment vs. control differential attrition	3	.025	.070	-0.055 -.068

Table 19

Characteristics of Included Studies – Feeding

Study	Study Design/ Control Type [a]	Total *n* (Tx *n*)	Child's Mean age	Eligibility [b]	Target Behavior	Intervention Name[d]	Tx length/ attrition[e]	Intervention Characteristics[f]
Aboud et al., 2009 Bangladesh	Cluster-Randomized Field Trial; EC	203 (108)	13.87 mos.	NC	Child self-feeding, Maternal responsiveness to child's mealtime cues	+Responsive Feeding Manual	18*/NP/NP OD: 4.4 TCD: -5.5%	B/SP/C/G/HB/B
Morawska et al., 2014 Australia	RCT; WLC	86 (44)	44.6 mos.	PC	Mealtime cooperation, consistent discipline/ routine	+Hassle Free Mealtimes Tripe P Group	1/120/120 OD: 13% TCD: 6.8%	CBT/IP/NC/G/U/NB
Skouteris et al., 2015 Australia	RCT; WLC	250 (104)	33 mos.	NC	Eating habits	+MEND (Mind, Exercise, Nutrition… Doit!) 2-4	10/90/900 OD: 16% TCD: 6.4%	CBT/UI/C/G/C/NB

Note. [a] *Study design:* WLC: Waitlist Control; EC: Educational Control; [b] *Eligibility* NC = no pre-identified concern, diagnosis, or clinical cutoff needed [d] +:intervention is manualized. [e] *Number of Sessions/Minutes per session/Dosage:* First number is number of sessions, next is how many minutes per session, and last is dosage (number of sessions x minutes); OD = Overall Dropout Rate in %; TCD = Treatment/Control Differential in % - negative numbers mean greater attrition in control group, no sign otherwise means greater attrition in treatment group. Formulas for both are in appendix. [f] *Intervention Characteristics Grouping (in order): Theoretical Orientation* (B = Behavioral components; CBT= Cognitive Behavioral components; *Type of prevention or intervention* (SP= selective Prevention; IP = Indicated Prevention; UI = Universal Intervention); *Child Involvement* (C = Child involved; NC = Child not involved); *Delivery format*: G=Had group elements; *Treatment Setting* (HB = Home Based, C = Community, U = University); *Booster* (B = Booster; NB = No Booster). Other Abbreviations: WLC: Waitlist Control; EC: Educational Control; NP = data not provided not able to calculate from data provided. *6 sessions on responsive feeding; rest on nutrition/child development.

Table 20

Effect Sizes of Included Studies – Feeding

Study	Outcome Measures	Hedges *g (SE)* and time reported	Long term effects reported?
Aboud et al., 2009 Bangladesh	Behavioral Observations of Child Feeding Behaviors practices during a mid-day meal	g= .689 (SE = .16) PI (2 weeks after sessions ended)	5 mos. across groups
Morawska et al., 2014 Australia	PAFTA CAPES	g=.429 (SE=.24) PI	6 mo. For intervention group only
Skouteris et al.., 2015 Australia	CEBQ, Child Food Neophobia Scale (Pilner)	g=.131 (SE = .16) PI	6 and 12 mos. across groups

Note. Abbreviations: CAPES - Child Adjustment and Parent Efficacy Scale; *PAFTA*; Child Feeding Behaviors – Parent Self Report; *CEBQ:* The Children's Eating Behavior Questionnaire; PI=post-intervention

Table 21

Descriptive statistics – Toileting Study

	k	*M*	*SD*	Range
Parent age (yrs.)	0	.	.	.
Child age (mos.)	1	27.45		
Percent female parent	0	.	.	.
Percent white parent	1	100		
Percent ≥ college education	0	.	.	.
Percent female child	1	31.43	.	.
Overall attrition	0	.	.	.
Treatment vs. control differential attrition	0	.	.	.

Table 22

Characteristics of Included Studies – Toileting

Study	Study Design/ Control Type [a]	Total *n*/ (Tx *n*)	Child's Mean age	Eligibility [b]	Target Behavior	Intervention Name	Tx length/ attrition [c]	Intervention Characteristics [d]
Vermandel et al.. 2008 Belgium	RCT/ None	39 WAD-T = 20 TP-T = 19	WAD-T: 26.1 mos. TP-T 28.8 mos.	NC	Toilet training	Wetting alarm diaper training (WAD-T) Timed Potty Training (TP-T)	Training session length NP, however investigators provided daily support calls for 5 days OD= .103 TCD =.01	For both groups: B/UI/C/I/HB/NB

Note. [a] *Study design:* RCT: Randomized controlled trial. [b] *Eligibility*: Eligibility: NC = no pre-identified concern, diagnosis, or clinical cutoff needed *Number of Sessions/Minutes per session/Dosage:* First number is number of sessions, next is how many minutes per session, and last is dosage (number of sessions x minutes); OD = Overall Dropout Rate in %; TCD = Treatment/Control Differential in % - negative numbers mean greater attrition in control group, no sign otherwise means greater attrition in treatment group. Formulas for both are in appendix. [d] *Intervention Characteristics Grouping (in order): Theoretical Orientation* (B = Behavioral components; *Type of prevention or intervention*: UI = Universal Intervention; *Child Involvement*: C = Child involved; *Delivery format*: I = Individual; ***Treatment Setting:*** HB = Home Based; *Booster Session?*: NB = No Booster; NP = data not provided not able to calculate from data provided.

Table 23

Effect Sizes of Included Studies – Toileting

Study	Outcome Measures	Hedges *g* (*SE*) and time reported	Long term effects reported?
Vermandel et al., 2008 Belgium	Number of toilet-trained children in the WAD-T and TPT method	.063 (.209) Post-intervention (5 days)	Yes, at 14 days and at 1 month

References

*indicates studies in meta-analysis and systematic review

*Abikoff, H. B., Thompson, M., Laver-Bradbury, C., Long, N., Forehand, R. L., Miller Brotman, L., Klein, R. G., Reiss, P., Huo, L., & Sonuga-Barke, E. (2015). Parent training for preschool ADHD: a randomized controlled trial of specialized and generic programs. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *56*(6), 618–631. https://doi.org/10.1111/jcpp.12346

*Aboud, F. E., Shafique, S., & Akhter, S. (2009). A responsive feeding intervention increases children's self-feeding and maternal responsiveness but not weight gain. *The Journal of Nutrition*, *139*(9), 1738–1743. https://doi.org/10.3945/jn.109.104885

*Adachi, Y., Sato, C., Nishino, N., Ohryoji, F., Hayama, J., & Yamagami, T. (2009). A brief parental education for shaping sleep habits in 4-month-old infants. *Clinical Medicine & Research*, *7*(3), 85–92. https://doi.org/10.3121/cmr.2009.814

Adamson, M., Morawska, A., & Sanders, M. R. (2013). Childhood feeding difficulties: a randomized controlled trial of a group-based parenting intervention. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *34*(5), 293–302. https://doi.org/10.1097/DBP.0b013e3182961a38

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Azrin, N. H., Sneed, T. J., & Foxx, R. M. (1974). Dry-bed training: rapid elimination of childhood enuresis. *Behaviour Research and Therapy*, *12*(3), 147–156. https://doi.org/10.1016/0005-7967(74)90111-9

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin*, *129*(2), 195–215. https://doi.org/10.1037/0033-2909.129.2.195

Barmish, A. J., & Kendall, P. C. (2005). Should parents be co-clients in cognitive-behavioral therapy for anxious youth? *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *34*(3), 569–581. https://doi.org/10.1207/s15374424jccp3403_12

Baumel, A., Mathur, N., Pawar, A., & Muench, F. (2021). Psychosocial interventions for children with externalized behavior problems: An updated meta-analysis of moderator effects. *Journal of Child and Family Studies*, *30*(1), 65–86. https://doi.org/10.1007/s10826-020-01863-6

Bayer, J. K., Beatson, R., Bretherton, L., Hiscock, H., Wake, M., Gilbertson, T., Mihalopoulos, C., Prendergast, L. A., & Rapee, R. M. (2018). Translational delivery of Cool Little Kids to prevent child internalising problems: Randomised controlled trial. *The Australian and New Zealand Journal of Psychiatry*, *52*(2), 181–191. https://doi.org/10.1177/0004867417726582

Bayer, J. K., Hastings, P. D., Sanson, A. V., Ukoumunne, O. C., & Rubin, K. H. (2010). Predicting mid-childhood internalising symptoms: A longitudinal community study. *The International Journal of Mental Health Promotion*, *12*(1), 5–17. https://doi.org/10.1080/14623730.2010.9721802

Beauchaine, T. P., Gartner, J., & Hagen, B. (2000). Comorbid depression and heart rate variability as predictors of aggressive and hyperactive symptom responsiveness during inpatient treatment of conduct-disordered, ADHD boys. *Aggressive Behavior*, *26*(6), 425–441. https://doi.org/10.1002/1098-2337(200011)26:6<425::aid-ab2>3.0.co;2-i

Beauchaine, T. P., Webster-Stratton, C., & Reid, M. J. (2005). Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: A latent growth curve analysis. *Journal of Consulting and Clinical Psychology*, *73*(3), 371–388. https://doi.org/10.1037/0022-006x.73.3.371

Birmaher, B., Ehmann, M., Axelson, D. A., Goldstein, B. I., Monk, K., Kalas, C., Kupfer, D., Gill, M. K., Leibenluft, E., Bridge, J., Guyer, A., Egger, H. L., & Brent, D. A. (2009). Schedule for affective disorders and schizophrenia for school-age children (K-SADS-PL) for the assessment of preschool children – A preliminary psychometric study. *Journal of Psychiatric Research*, *43*(7), 680–686. https://doi.org/10.1016/j.jpsychires.2008.10.003

Blum, N. J., Taubman, B., & Nemeth, N. (2003). Relationship between age at initiation of toilet training and duration of training: a prospective study. *Pediatrics*, *111*(4 Pt 1), 810–814. https://doi.org/10.1542/peds.111.4.810

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Wiley-Blackwell.

*Bradley, S. J., Jadaa, D.-A., Brody, J., Landy, S., Tallett, S. E., Watson, W., Shea, B., & Stephens, D. (2003). Brief psychoeducational parenting program: An evaluation and 1-year follow-up. *Journal of the American Academy of Child and Adolescent*

Psychiatry, *42*(10), 1171–1178. https://doi.org/10.1097/00004583-200310000-00007

*Brassart, E., Schelstraete, M.-A., & Roskam, I. (2017). What are the effects of a parent-implemented verbal responsive intervention on preschoolers with externalizing behavior problems? *Child & Family Behavior Therapy*, *39*(2), 108–130. https://doi.org/10.1080/07317107.2017.1307679

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological test administration by videoconference: A systematic review and meta-analysis. *Neuropsychology Review*, *27*, 174–186.

Briggs-Gowan, M. J., Carter, A. S., Bosson-Heenan, J., Guyer, A. E., & Horwitz, S. M. (2006). Are infant-toddler social-emotional and behavioral problems transient? *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(7), 849–858. https://doi.org/10.1097/01.chi.0000220849.48650.59

*Brotman, L. M., Klein, R. G., Kamboukos, D., Brown, E. J., Coard, S. I., & Sosinsky, L. S. (2003). Preventive intervention for urban, low-income preschoolers at familial risk for conduct problems: A randomized pilot study. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *32*(2), 246–257. https://doi.org/10.1207/s15374424jccp3202_10

Bufferd, S. J., Dougherty, L. R., Carlson, G. A., & Klein, D. N. (2011). Parent-reported mental health in preschoolers: findings using a diagnostic interview.

Comprehensive Psychiatry, *52*(4), 359–369. https://doi.org/10.1016/j.comppsych.2010.08.006

Card, N. A. (2015). *Applied meta-analysis for social science research*. Guilford Publications.

Carruth, B. R., Ziegler, P. J., Gordon, A., & Barr, S. I. (2004). Prevalence of picky eaters among infants and toddlers and their caregivers' decisions about offering a new food. *Journal of the American Dietetic Association*, *104*(1 Suppl 1), s57-64. https://doi.org/10.1016/j.jada.2003.10.024

*Chronis-Tuscano, A., Rubin, K. H., O'Brien, K. A., Coplan, R. J., Thomas, S. R., Dougherty, L. R., Cheah, C. S. L., Watts, K., Heverly-Fitt, S., Huggins, S. L., Menzer, M., Begle, A. S., & Wimsatt, M. (2015). Preliminary evaluation of a multimodal early intervention program for behaviorally inhibited preschoolers. *Journal of Consulting and Clinical Psychology*, *83*(3), 534–540. https://doi.org/10.1037/a0039043

Cicchetti, D., & Sroufe, L. A. (2000). The past as prologue to the future: the times, they've been a-changin'. *Development and Psychopathology*, *12*(3), 255–264. https://doi.org/10.1017/s0954579400003011

Comer, J. S., Chow, C., Chan, P. T., Cooper-Vince, C., & Wilson, L. A. S. (2013). Psychosocial treatment efficacy for disruptive behavior problems in very young children: A meta-analytic examination. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(1), 26–36. https://doi.org/10.1016/j.jaac.2012.10.001

Comer, J. S., Hong, N., Poznanski, B., Silva, K., & Wilson, M. (2019). Evidence Base Update on the treatment of early childhood anxiety and related problems. *Journal*

of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53, *48*(1), 1–15. https://doi.org/10.1080/15374416.2018.1534208

Coplan, R. J., Schneider, B. H., Matheson, A., & Graham, A. (2010). 'Play skills' for shy children: development of a Social Skills Facilitated Playearly intervention program for extremely inhibited preschoolers. *Infant and Child Development*. https://doi.org/10.1002/icd.668

Daniels, L. A., Magarey, A., Battistutta, D., Nicholson, J. M., Farrell, A., Davidson, G., & Cleghorn, G. (2009). The NOURISH randomised control trial: positive feeding practices and food preferences in early childhood - a primary prevention program for childhood obesity. *BMC Public Health*, *9*(1), 387. https://doi.org/10.1186/1471-2458-9-387

Dasari, M., & Knell, S. M. (2015). Cognitive-behavioral play therapy for children with anxiety and phobias. In H. Kaduson (Ed.), *Short-term play therapy for children* (pp. 25–52).

de Carvalho Mrad, F. C., da Silva, M. E., Moreira Lima, E., Bessa, A. L., de Bessa Junior, J., Netto, J. M. B., & de Almeida Vasconcelos, M. M. (2021). Toilet training methods in children with normal neuropsychomotor development: A systematic review. *Journal of Pediatric Urology*, *17*(5), 635–643. https://doi.org/10.1016/j.jpurol.2021.05.010

Deke, J., & Chiang, H. (2017). The WWC attrition standard: Sensitivity to assumptions and opportunities for refining and adapting to new contexts, *Evaluation Review*, *41*(2), 130–154. https://doi.org/10.1177/0193841X16670047

*Dishion, T. J., Shaw, D., Connell, A., Gardner, F., Weaver, C., & Wilson, M. (2008). The family check-up with high-risk indigent families: Preventing problem behavior by increasing parents' positive behavior support in early childhood. *Child Development*, *79*(5), 1395–1414. https://doi.org/10.1111/j.1467-8624.2008.01195.x

*Dittman, C. K., Farruggia, S. P., Keown, L. J., & Sanders, M. R. (2016). Dealing with disobedience: An evaluation of a brief parenting intervention for young children showing noncompliant behavior problems. *Child Psychiatry and Human Development*, *47*(1), 102–112. https://doi.org/10.1007/s10578-015-0548-9

Dodd, H. F., Hudson, J. L., & Rapee, R. M. (2017). Temperament in youth internalizing disorders. In *Treatments for Psychological Problems and Syndromes* (pp. 504–524). John Wiley & Sons, Ltd.

*Eckerberg, B. (2007). Treatment of sleep problems in families with small children: is written information enough? *Acta Paediatrica (Oslo, Norway: 1992)*, *91*(8), 952–959. https://doi.org/10.1111/j.1651-2227.2002.tb02861.x

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(3–4), 313–337. https://doi.org/10.1111/j.1469-7610.2006.01618.x

Egger, H. L., Erkanli, A., Keeler, G., Potts, E., Walter, B. K., & Angold, A. (2006). Test-retest reliability of the Preschool Age Psychiatric Assessment (PAPA). *Journal of*

the American Academy of Child and Adolescent Psychiatry, *45*(5), 538–549. https://doi.org/10.1097/01.chi.0000205705.71194.b8

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *37*(1), 215–237. https://doi.org/10.1080/15374410701820117

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *The British Journal of Mathematical and Statistical Psychology*, *63*(Pt 3), 665–694. https://doi.org/10.1348/000711010X502733

Ginsburg, G. S., & Schlossberg, M. C. (2002). Family-based treatment of childhood anxiety disorders. *International Review of Psychiatry (Abingdon, England)*, *14*(2), 143–154. https://doi.org/10.1080/09540260220132662

*Hall, W. A., Hutton, E., Brant, R. F., Collet, J. P., Gregg, K., Saunders, R., Ipsiroglu, O., Gafni, A., Triolet, K., Tse, L., Bhagat, R., & Wooldridge, J. (2015). A randomized controlled trial of an intervention for infants' behavioral sleep problems. *BMC Pediatrics*, *15*(1). https://doi.org/10.1186/s12887-015-0492-7

Harris, M., Andrews, K., Gonzalez, A., Prime, H., & Atkinson, L. (2020). Technology-assisted parenting interventions for families experiencing social disadvantage: A meta-analysis. *Prevention Science: The Official Journal of the Society for Prevention Research*, *21*(5), 714–727. https://doi.org/10.1007/s11121-020-01128-0

Harwood, J., Fernández, L., Vallejo, V., & Day, C. (2022). Baby and us: Community-based, Feasibility Trial of a Psychosocial Intervention for New Parents and their

Infants. *Journal of Prevention (2022)*, *43*(5), 589–604. https://doi.org/10.1007/s10935-022-00685-0

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Helle, C., Hillesund, E. R., Omholt, M. L., & Øverby, N. C. (2017). Early food for future health: a randomized controlled trial evaluating the effect of an eHealth intervention aiming to promote healthy food habits from early childhood. *BMC Public Health*, *17*(1). https://doi.org/10.1186/s12889-017-4731-8

Hemmi, M. H., Wolke, D., & Schneider, S. (2011). Associations between problems with crying, sleeping and/or feeding in infancy and long-term behavioural outcomes in childhood: a meta-analysis. *Archives of Disease in Childhood*, *96*(7), 622–629. https://doi.org/10.1136/adc.2010.191312

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ (Clinical Research Ed.)*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Higgins, J. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2022). *Cochrane Handbook for Systematic Reviews of Interventions version* (Vol. 6).

Holland, M. L., Malmberg, J., & Peacock, G. G. (2017). *Emotional and behavioral problems of young children: Effective interventions in the preschool and kindergarten years*. Guilford Publications.

Hudson, J. L., Minihan, S., Chen, W., Carl, T., Fu, M., Tully, L., Kangas, M., Rosewell, L., McDermott, E. A., Wang, Y., Stubbs, T., & Martiniuk, A. (2023). Interventions

for young children's mental health: A review of reviews. *Clinical Child and Family Psychology Review*, *26*(3), 593–641. https://doi.org/10.1007/s10567-023-00443-6

*Hutchings, J., Bywater, T., Daley, D., Gardner, F., Whitaker, C., Jones, K., Eames, C., & Edwards, R. T. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *BMJ*, *334*(7595), 678. https://doi.org/10.1136/bmj.39126.620799.55

IBM Corp. (2020). IBM SPSS Statistics for Mac (Version 27.0) [Computer software].

Institute of Education Sciences (n.d.) Attrition. [PowerPoint Slides]. https://ies.ed.gov/ncee/wwc/Docs/OnlineTraining/wwc_training_m2.pdf

JASP Team (2022). JASP (Version 0.16.1)[Computer software].

Jeong, J., Franchett, E. E., Ramos de Oliveira, C. V., Rehmani, K., & Yousafzai, A. K. (2021). Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLoS Medicine*, *18*(5), e1003602. https://doi.org/10.1371/journal.pmed.1003602

Jewell, C., Wittkowski, A., & Pratt, D. (2022). The impact of parent-only interventions on child anxiety: A systematic review and meta-analysis. *Journal of Affective Disorders*, *309*, 324–349. https://doi.org/10.1016/j.jad.2022.04.082

Kahn, M., Livne-Karp, E., Juda-Hanael, M., Omer, H., Tikotzky, L., Anders, T. F., & Sadeh, A. (2020). Behavioral interventions for infant sleep problems: the role of parental cry tolerance and sleep-related cognitions. *Journal of Clinical Sleep Medicine: JCSM: Official Publication of the American Academy of Sleep Medicine*, *16*(8), 1275–1283. https://doi.org/10.5664/jcsm.8488

Kaminski, J. W., & Claussen, A. H. (2017). Evidence base update for psychosocial treatments for disruptive behaviors in children. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *46*(4), 477–499. https://doi.org/10.1080/15374416.2017.1310044

*Kennedy, S. J., Rapee, R. M., & Edwards, S. L. (2009). A selective intervention program for inhibited preschool-aged children of parents with an anxiety disorder: Effects on current anxiety disorders and temperament. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(6), 602–609. https://doi.org/10.1097/chi.0b013e31819f6fa9

Lavigne, J. V., Lebailly, S. A., Hopkins, J., Gouze, K. R., & Binns, H. J. (2009). The prevalence of ADHD, ODD, depression, and anxiety in a community sample of 4-year-olds. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *38*(3), 315–328. https://doi.org/10.1080/15374410902851382

Lebowitz, E. R., Marin, C., Martino, A., Shimshoni, Y., & Silverman, W. K. (2020). Parent-based treatment as efficacious as cognitive-behavioral therapy for Childhood Anxiety: A randomized noninferiority study of supportive parenting for Anxious Childhood Emotions. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(3), 362–372. https://doi.org/10.1016/j.jaac.2019.02.014

Lenze, S. N., Pautsch, J., & Luby, J. (2011). Parent-child interaction therapy emotion development: a novel treatment for depression in preschool children. *Depression and Anxiety*, *28*(2), 153–159. https://doi.org/10.1002/da.20770

Lin, Y.W., & Bratton, S. C. (2015). A meta-analytic review of child-centered play therapy approaches. *Journal of Counseling and Development: JCD*, *93*(1), 45–58. https://doi.org/10.1002/j.1556-6676.2015.00180.x

Luby, J. L., Barch, D. M., Whalen, D., Tillman, R., & Freedland, K. E. (2018). A randomized controlled trial of parent-child psychotherapy targeting emotion development for early childhood depression. *The American Journal of Psychiatry*, *175*(11), 1102–1110. https://doi.org/10.1176/appi.ajp.2018.18030321

Luby, J. L., Heffelfinger, A. K., Mrakotsky, C., Brown, K. M., Hessler, M. J., Wallis, J. M., & Spitznagel, E. L. (2003). The clinical picture of depression in preschool children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(3), 340–348. https://doi.org/10.1097/00004583-200303000-00015

Luby, J., Lenze, S., & Tillman, R. (2012). A novel early intervention for preschool depression: findings from a pilot randomized controlled trial: RCT PCIT-ED preschool MDD. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *53*(3), 313–322. https://doi.org/10.1111/j.1469-7610.2011.02483.x

Lukens, C. T., & Silverman, A. H. (2014). Systematic review of psychological interventions for pediatric feeding problems. *Journal of Pediatric Psychology*, *39*(8), 903–917. https://doi.org/10.1093/jpepsy/jsu040

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2006). A meta-analysis of parent training: moderators and follow-up effects. *Clinical Psychology Review*, *26*(1), 86–104. https://doi.org/10.1016/j.cpr.2005.07.004

*MacKenzie, E. P., & Hilgedick, J. M. (2000). The computer-assisted parenting program (CAPP): The use of a computerized behavioral parent training program as an educational tool. *Child & Family Behavior Therapy*, *21*(4), 23–43. https://doi.org/10.1300/j019v21n04_02

*Markie-Dadds, C., & Sanders, M. R. (2006). Self-directed Triple P (Positive Parenting Program) for mothers with children at-risk of developing conduct problems. *Behavioural and Cognitive Psychotherapy*, *34*(3), 259–275. https://doi.org/10.1017/s1352465806002797

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, *22*(3), 491–495. https://doi.org/10.1017/s0954579410000222

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review*, *34*(3), 267–286. https://doi.org/10.1080/02796015.2005.12086287

McLuckie, A., Landers, A. L., Curran, J. A., Cann, R., Carrese, D. H., Nolan, A., Corrigan, K., & Carrey, N. J. (2019). A scoping review of mental health prevention and intervention initiatives for infants and preschoolers at risk for socio-emotional difficulties. *Systematic Reviews*, *8*(1), 183. https://doi.org/10.1186/s13643-019-1043-3

Meltzer, L. J., & Mindell, J. A. (2006). Sleep and sleep disorders in children and adolescents. *The Psychiatric Clinics of North America*, *29*(4), 1059–1076. https://doi.org/10.1016/j.psc.2006.08.004

Meltzer, L. J., & Mindell, J. A. (2014). Systematic review and meta-analysis of behavioral interventions for pediatric insomnia. *Journal of Pediatric Psychology*, *39*(8), 932–948. https://doi.org/10.1093/jpepsy/jsu041

Meltzer, L. J., Wainer, A., Engstrom, E., Pepa, L., & Mindell, J. A. (2021a). Seeing the Whole Elephant: a scoping review of behavioral treatments for pediatric insomnia. *Sleep Medicine Reviews*, *56*(101410), 101410. https://doi.org/10.1016/j.smrv.2020.101410

Meltzer, L. J., Williamson, A. A., & Mindell, J. A. (2021b). Pediatric sleep health: It matters, and so does how we define it. *Sleep Medicine Reviews*, *57*(101425), 101425. https://doi.org/10.1016/j.smrv.2021.101425

Mihelic, M., Morawska, A., & Filus, A. (2017). Effects of early parenting interventions on parents and infants: A meta-analytic review. *Journal of Child and Family Studies*, *26*(6), 1507–1526. https://doi.org/10.1007/s10826-017-0675-y

*Mindell, J. A., Du Mond, C. E., Sadeh, A., Telofski, L. S., Kulkarni, N., & Gunn, E. (2011). Efficacy of an internet-based intervention for infant and toddler sleep disturbances. *Sleep*, *34*(4), 451–458. https://doi.org/10.1093/sleep/34.4.451

Mindell, J. A., Kuhn, B., Lewin, D. S., Meltzer, L. J., Sadeh, A., & American Academy of Sleep Medicine (2006). Behavioral treatment of bedtime problems and night wakings in infants and young children. *Sleep*, *29*(10), 1263–1276. https://doi.org/10.1093/sleep/29.10.1263

*Mindell, J. A., Telofski, L. S., Wiegand, B., & Kurtz, E. S. (2009). A nightly bedtime routine: Impact on sleep in young children and maternal mood. *Sleep*, *32*(5), 599–606. https://doi.org/10.1093/sleep/32.5.599

Mingebach, T., Kamp-Becker, I., Christiansen, H., & Weber, L. (2018). Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PloS One*, *13*(9), e0202855. https://doi.org/10.1371/journal.pone.0202855

*Morawska, A., Adamson, M., Hinchliffe, K., & Adams, T. (2014). Hassle Free Mealtimes Triple P: A randomised controlled trial of a brief parenting group for childhood mealtime difficulties. *Behaviour Research and Therapy*, *53*, 1–9. https://doi.org/10.1016/j.brat.2013.11.007

*Morawska, A., Haslam, D., Milne, D., & Sanders, M. R. (2011). Evaluation of a brief parenting discussion group for parents of young children. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *32*(2), 136–145. https://doi.org/10.1097/dbp.0b013e3181f17a28

*Morawska, A., & Sanders, M. R. (2006). Self-administered behavioral family intervention for parents of toddlers: Part I. Efficacy. *Journal of Consulting and Clinical Psychology*, *74*(1), 10–19. https://doi.org/10.1037/0022-006x.74.1.10

*Nicholson, B., Anderson, M., Fox, R., & Brenner, V. (2002). One family at a time: A prevention program for at-risk parents. *Journal of Counseling and Development: JCD*, *80*(3), 362–371. https://doi.org/10.1002/j.1556-6678.2002.tb00201.x

*Nixon, R. D. V. (2001). Changes in hyperactivity and temperament in behaviourally disturbed preschoolers after parent–child interaction therapy (PCIT). *Behaviour*

Change: Journal of the Australian Behaviour Modification Association, *18*(3), 168–176. https://doi.org/10.1375/bech.18.3.168

*Nixon, R. D. V., Sweeney, L., Erickson, D. B., & Touyz, S. W. (2003). Parent-child interaction therapy: a comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *Journal of Consulting and Clinical Psychology*, *71*(2), 251–260. https://doi.org/10.1037/0022-006x.71.2.251

Nock, M. K. (2003). Progress review of the psychosocial treatment of child conduct problems. *Clinical Psychology: A Publication of the Division of Clinical Psychology of the American Psychological Association*, *10*(1), 1–28. https://doi.org/10.1093/clipsy/10.1.1

Ooi, J., Dodd, H. F., Meiser-Stedman, R., Hudson, J. L., Bridges, J., & Pass, L. (2022). The efficacy of interventions for behaviourally inhibited preschool-aged children: A meta-analysis. *Journal of Anxiety Disorders*, *88*(102559), 102559. https://doi.org/10.1016/j.janxdis.2022.102559

Owens, E. B., Hinshaw, S. P., Kraemer, H. C., Arnold, L. E., Abikoff, H. B., Cantwell, D. P., Conners, C. K., Elliott, G., Greenhill, L. L., Hechtman, L., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Pelham, W. E., Severe, J. B., Swanson, J. M., Vitiello, B., Wells, K. C., & Wigal, T. (2003). Which treatment for whom for ADHD? Moderators of treatment response in the MTA. *Journal of Consulting and Clinical Psychology*, *71*(3), 540–552. https://doi.org/10.1037/0022-006x.71.3.540

Park, J., Kim, S. Y., & Lee, K. (2022). Effectiveness of behavioral sleep interventions on children's and mothers' sleep quality and maternal depression: a systematic review

and meta-analysis. *Scientific Reports*, *12*(1), 4172. https://doi.org/10.1038/s41598-022-07762-8

Patterson, G. R. (1974). Interventions for boys with conduct problems: multiple settings, treatments, and criteria. *Journal of Consulting and Clinical Psychology*, *42*(4), 471–481. https://doi.org/10.1037/h0036731

*Paul, I. M., Savage, J. S., Anzman-Frasca, S., Marini, M. E., Mindell, J. A., & Birch, L. L. (2016). INSIGHT responsive parenting intervention and infant sleep. *Pediatrics*, *138*(1). https://doi.org/10.1542/peds.2016-0762

Paulus, F. W., Backes, A., Sander, C. S., Weber, M., & von Gontard, A. (2015). Anxiety disorders and behavioral inhibition in preschool children: A population-based study. *Child Psychiatry and Human Development*, *46*(1), 150–157. https://doi.org/10.1007/s10578-014-0460-8

*Perrin, E. C., Sheldrick, R. C., McMenamy, J. M., Henson, B. S., & Carter, A. S. (2014). Improving parenting skills for families of young children in pediatric settings: A randomized clinical trial. *JAMA Pediatrics*, *168*(1), 16. https://doi.org/10.1001/jamapediatrics.2013.2919

Pinquart, M., & Teubert, D. (2010). Effects of parenting education with expectant and new parents: a meta-analysis. *Journal of Family Psychology: JFP: Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *24*(3), 316–327. https://doi.org/10.1037/a0019691

*Posthumus, J. A., Raaijmakers, M. A. J., Maassen, G. H., van Engeland, H., & Matthys, W. (2012). Sustained effects of incredible years as a preventive intervention in

preschool children with conduct problems. *Journal of Abnormal Child Psychology*, *40*(4), 487–500. https://doi.org/10.1007/s10802-011-9580-9

Powell, D., Dunlap, G., & Fox, L. (2006). Prevention and intervention for the challenging behaviors of toddlers and preschoolers. *Infants and Young Children*, *19*(1), 25–35. https://doi.org/10.1097/00001163-200601000-00004

Rapee, R. M. (2013). The preventative effects of a brief, early intervention for preschool-aged children at risk for internalising: follow-up into middle adolescence. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *54*(7), 780–788. https://doi.org/10.1111/jcpp.12048

Rapee, R. M., & Coplan, R. J. (2010). Conceptual relations between anxiety disorder and fearful temperament. *New Directions for Child and Adolescent Development*, *2010*(127), 17–31. https://doi.org/10.1002/cd.260

*Rapee, R. M., Kennedy, S. J., Ingram, M., Edwards, S. L., & Sweeney, L. (2010). Altering the trajectory of anxiety in at-risk young children. *The American Journal of Psychiatry*, *167*(12), 1518–1525. https://doi.org/10.1176/appi.ajp.2010.09111619

*Reid, G. J., Stewart, M., Vingilis, E., Dozois, D. J. A., Wetmore, S., Jordan, J., Dickie, G., Osmun, W. E., Wade, T. J., Brown, J. B., & Zaric, G. S. (2013). Randomized trial of distance-based treatment for young children with discipline problems seen in primary health care. *Family Practice*, *30*(1), 14–24. https://doi.org/10.1093/fampra/cms051

*Reid, M. J., Walter, A. L., & O'Leary, S. G. (1999). Treatment of young children's bedtime refusal and nighttime wakings: a comparison of "standard" and graduated

ignoring procedures. *Journal of Abnormal Child Psychology*, *27*(1), 5–16. https://doi.org/10.1023/a:1022606206076

Reuter, A., Silfverdal, S.-A., Lindblom, K., & Hjern, A. (2020). A systematic review of prevention and treatment of infant behavioural sleep problems. *Acta Paediatrica (Oslo, Norway: 1992)*, *109*(9), 1717–1732. https://doi.org/10.1111/apa.15182

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems - a meta-analytic review. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(1), 99–111. https://doi.org/10.1111/j.1469-7610.2005.01544.x

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review*, *32*(4), 251–262. https://doi.org/10.1016/j.cpr.2012.01.005

*Sanders, M. R., Markie-Dadds, C., Tully, L. A., & Bor, W. (2000). The triple P-positive parenting program: a comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *Journal of Consulting and Clinical Psychology*, *68*(4), 624–640. https://doi.org/10.1037//0022-006x.68.4.624

Sarkadi, A., Sampaio, F., Kelly, M. P., & Feldman, I. (2014). A novel approach used outcome distribution curves to estimate the population-level impact of a public health intervention. *Journal of Clinical Epidemiology*, *67*(7), 785–792. https://doi.org/10.1016/j.jclinepi.2013.12.012

Satter, E. (1990). The feeding relationship: Problems and interventions. *The Journal of Pediatrics*, *117*(2), S181–S189. https://doi.org/10.1016/s0022-3476(05)80017-4

Savage, J. S., Hohman, E. E., Marini, M. E., Shelly, A., Paul, I. M., & Birch, L. L. (2018). INSIGHT responsive parenting intervention and infant feeding practices: randomized clinical trial. *The International Journal of Behavioral Nutrition and Physical Activity*, *15*(1). https://doi.org/10.1186/s12966-018-0700-6

*Schlarb, A. A., Brandhorst, I., Peters, E., & Hautzinger, M. (2017). Telephone support in an internet-based treatment for sleep problems in early childhood. *Journal of Psychiatry*, *21*(1). https://doi.org/10.4172/2378-5756.1000435

*Shaw, D. S., Dishion, T. J., Supplee, L., Gardner, F., & Arnds, K. (2006). Randomized trial of a family-centered approach to the prevention of early conduct problems: 2-year effects of the family check-up in early childhood. *Journal of Consulting and Clinical Psychology*, *74*(1), 1–9. https://doi.org/10.1037/0022-006X.74.1.1

Shi, L., Chu, H., & Lin, L. (2020). A Bayesian approach to assessing small-study effects in meta-analysis of a binary outcome with controlled false positive rate. *Research Synthesis Methods*, *11*(4), 535–552. https://doi.org/10.1002/jrsm.1415

Silverman, W. K., Pina, A. A., & Viswesvaran, C. (2008). Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *37*(1), 105–130. https://doi.org/10.1080/15374410701817907

*Skouteris, H., Hill, B., McCabe, M., Swinburn, B., & Busija, L. (2016). A parent-based intervention to promote healthy eating and active behaviours in pre-school children: evaluation of the MEND 2-4 randomized controlled trial: Promoting healthy

behaviours in pre-school children. *Pediatric Obesity*, *11*(1), 4–10. https://doi.org/10.1111/ijpo.12011

Sleed, M., Li, E. T., Vainieri, I., & Midgley, N. (2023). The evidence-base for psychodynamic interventions with children under 5 years of age and their caregivers: A systematic review and meta-analysis. *Journal of Infant, Child, and Adolescent Psychotherapy: JICAP*, *22*(3), 179–214. https://doi.org/10.1080/15289168.2023.2223739

*Somech, L. Y., & Elizur, Y. (2012). Promoting self-regulation and cooperation in pre-kindergarten children with conduct problems: a randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(4), 412–422. https://doi.org/10.1016/j.jaac.2012.01.019

*Sonuga-Barke, E. J., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: a randomized, controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(4), 402–408. https://doi.org/10.1097/00004583-200104000-00008

*Sourander, A., McGrath, P. J., Ristkari, T., Cunningham, C., Huttunen, J., Lingley-Pottie, P., Hinkka-Yli-Salomäki, S., Kinnunen, M., Vuorio, J., Sinokki, A., Fossum, S., & Unruh, A. (2016). Internet-assisted parent training intervention for disruptive behavior in 4-year-old children: A randomized clinical trial. *JAMA Psychiatry (Chicago, Ill.)*, *73*(4), 378. https://doi.org/10.1001/jamapsychiatry.2015.3411

Southam-Gerow, M. A., & Prinstein, M. J. (2014). Evidence base updates: The evolution of the evaluation of psychological treatments for children and adolescents. *Journal*

of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53, *43*(1), 1–6. https://doi.org/10.1080/15374416.2013.855128

*St James-Roberts, I., & Gillham, P. (2001). Use of a behavioural programme in the first 3 months to prevent infant crying and sleeping problems. *Journal of Paediatrics and Child Health*, *37*(3), 289–297. https://doi.org/10.1046/j.1440-1754.2001.00699.x

*Stevens, J., Splaingard, D., Webster-Cheng, S., Rausch, J., & Splaingard, M. (2019). A randomized trial of a self-administered parenting intervention for infant and toddler insomnia. *Clinical Pediatrics*, *58*(6), 633–640. https://doi.org/10.1177/0009922819832030

*Stremler, R., Hodnett, E., Kenton, L., Lee, K., Weiss, S., Weston, J., & Willan, A. (2013). Effect of behavioural-educational intervention on sleep for primiparous women and their infants in early postpartum: multisite randomised controlled trial. *BMJ (Clinical Research Ed.)*, *346*(mar20 1), f1164–f1164. https://doi.org/10.1136/bmj.f1164

Tanner, K., Dempster, R., Castillo, A., Burdo-Hartman, W., Halpin, E., Rausch, J., & Stevens, J. (2022). Randomized trial of a self-administered parenting intervention for selective eating in young children. *Eating Behaviors*, *46*(101646), 101646. https://doi.org/10.1016/j.eatbeh.2022.101646

Tarver, J., Daley, D., Lockwood, J., & Sayal, K. (2014). Are self-directed parenting interventions sufficient for externalising behaviour problems in childhood? A

systematic review and meta-analysis. *European Child & Adolescent Psychiatry*, *23*(12), 1123–1137. https://doi.org/10.1007/s00787-014-0556-5

Taylor, C. M., Wernimont, S. M., Northstone, K., & Emmett, P. M. (2015). Picky/fussy eating in children: Review of definitions, assessment, prevalence and dietary intakes. *Appetite*, *95*, 349–359. https://doi.org/10.1016/j.appet.2015.07.026

Thongseiratch, T., Leijten, P., & Melendez-Torres, G. J. (2020). Online parent programs for children's behavioral problems: a meta-analytic review. *European Child & Adolescent Psychiatry*, *29*(11), 1555–1568. https://doi.org/10.1007/s00787-020-01472-0

Tremblay, R. E. (2004). Decade of Behavior Distinguished Lecture: Development of physical aggression during infancy. *Infant Mental Health Journal*, *25*(5), 399–407. https://doi.org/10.1002/imhj.20015

*Van Zeijl, J., Mesman, J., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Juffer, F., Stolk, M. N., Koot, H. M., & Alink, L. R. A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of 1- to 3-year-old children at risk for externalizing behavior problems: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *74*(6), 994–1005. https://doi.org/10.1037/0022-006x.74.6.994

*Vermandel, A., Weyler, J., De Wachter, S., & Wyndaele, J.-J. (2008). Toilet training of healthy young toddlers: a randomized trial between a daytime wetting alarm and timed potty training. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *29*(3), 191–196. https://doi.org/10.1097/DBP.0b013e31816c433a

Viechtbauer, W. (2010). Conducting meta-analyses in *R*with the metafor Package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Wardle, J., Carnell, S., & Cooke, L. (2005). Parental control over feeding and children's fruit and vegetable intake: how are they related? *Journal of the American Dietetic Association*, *105*(2), 227–232. https://doi.org/10.1016/j.jada.2004.11.006

Weber, L., Kamp-Becker, I., Christiansen, H., & Mingebach, T. (2019). Treatment of child externalizing behavior problems: a comprehensive review and meta-meta-analysis on effects of parent-based interventions on parental characteristics. *European Child & Adolescent Psychiatry*, *28*(8), 1025–1036. https://doi.org/10.1007/s00787-018-1175-3

Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., Jensen-Doss, A., Hawley, K. M., Krumholz Marchette, L. S., Chu, B. C., Weersing, V. R., & Fordwood, S. R. (2017). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *The American Psychologist*, *72*(2), 79–117. https://doi.org/10.1037/a0040360

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychological Bulleti n*, *132*(1), 132–149. https://doi.org/10.1037/0033-2909.132.1.132

Whalen, D. J., Sylvester, C. M., & Luby, J. L. (2017). Depression and anxiety in preschoolers. *Child and Adolescent Psychiatric Clinics of North America*, *26*(3), 503–522. https://doi.org/10.1016/j.chc.2017.02.006

What Works Clearinghouse Procedures and Standards Handbook, Version 5.0; U.S. Department of Education, Institute of Education Sciences. (2022). *National Center for Education Evaluation and Regional Assistance*. This report Is Available on the What Works Clearinghouse Website at https://ies.ed.gov/ncee/wwc/Handbooks

What Works Clearinghouse. (2014). *Assessing Attrition Bias* [White Paper]. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_attrition_v2.1.pdf

Wichstrøm, L., Berg-Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E., & Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers: Psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *53*(6), 695–705. https://doi.org/10.1111/j.1469-7610.2011.02514.x

Vita

Name	Hara Stephanou
Baccalaureate Degree	Bachelor of Arts, Adelphi University Garden City, New York Major: Psychology
Date Graduated	May 2011
Other Degrees and Certificates	Master of Science, St. John's University Jamaica, New York Major: School Psychology
Date Graduated	September 2018